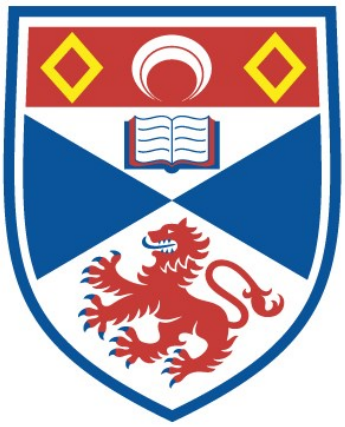
THEORETICAL STUDIES OF ENZYME INHIBITION

Colin Michael Edge

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews



1989

THEORETICAL STUDIES OF ENZYME INHIBITION

A Thesis

presented for the degree of

Doctor of Philosophy

in the Faculty of Science of the

University of St Andrews

by

Colin Michael Edge, B.Sc.

St Leonard's College, 1986

I, Colin Michael Edge, hereby certify that this thesis has been composed by myself, that it is a record of my own work, and that it has not been accepted in partial or complete fulfilment of any other degree or professional qualification.

Signed

Date December 1st 1986

I was admitted to the Faculty of Science of the University of St Andrews under Ordinance General No 12 on October 1st 1983 and as a candidate for the degree of Ph.D. on February 1985.

Signed Date December 1st 1986

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate to the Degree of Ph.D.

Signature of Supervisor Date December 1st 1986

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker.

Signed Date December 1st 1986

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr Colin Thomson for providing me with the opportunity to study this intriguing enzyme and for all his encouragement, suggestions and advice whilst I was a research student.

I have also benefitted from detailed discussions with a number of people over this period: these are Dr Chris Reynolds, Mr Derek Higgins and especially Dr John Ball, who not only had the patience to answer all my questions, but even provided me with a place to stay in the latter days of this study.

The National Foundation for Cancer Research must also be thanked for providing me with a research studentship and for providing financial support in order for me to attend the Faraday Symposium in 1984 at Cambridge and the ChemGraf Users Group Meeting in 1984 at Oxford. The NFCR also donated the necessary funds for the Tektronix graphics system, which played such an important part in this work. As the NFCR is a charitable organisation, I must thank the thousands of anonymous contributors to this cause.

I would also like to thank Prof. J. Pople for copies of the GAUSSIAN 80 and GAUSSIAN 82 programs, and Andy Micklethwaite, a Systems Analyst at Tektronix, who provided me with a simple graphics program which stimulated my burgeoning interest in computer graphics.

A special acknowledgement is due to the St Andrews University Computing Laboratory for the computing facilities, to the advisors for their untiring efforts in solving problems and especially to the operators for their cheerful service.

I would like to thank all my friends and family who have kept me going over the last few years, especially my brother Jonathan, who drew some of the diagrams because he wanted to be mentioned in the Acknowledgements.

Last of all, I must thank my parents for their support and encouragement.

ABSTRACT

The glyoxalase enzyme system catalyses the conversion of methylglyoxal to D-lactic acid. The first of the two component enzymes, glyoxalase I, is responsible for the transfer of two protons in an isomerisation reaction. This enzyme has been ascribed a rôle in tumorigenesis in the past and some of its inhibitors are known to be carcinostatic.

This thesis describes quantum chemical calculations on the enzyme mechanism and on some enzyme inhibitors.

The calculations on the mechanism of the enzyme take the form of studies of model reaction schemes, with minimal and split-valence basis sets. The calculation of the energies of various intermediates has led to the evaluation of different pathways as models of the enzyme mechanism. The comparison of different substituted compounds has led to further conclusions on the part played by the sulphur atom in the enzyme-catalysed reaction.

Two main groups of inhibitor molecules are discussed; these are flavone and coumarin derivatives. The molecular electrostatic potential of these molecules has been calculated on various surfaces, using a minimal basis set, to attempt to correlate this property with the compounds' inhibitory power.

A FORTRAN program is presented which depicts calculated properties on the surfaces. This program allowed the identification of various regions which seemed to be indicative of the inhibitory strength of the compounds.

CONTENTS

CHAPTER 1

INTRODUCTION

Cancer is perhaps the most feared disease in the Western World. Whilst many other diseases, especially the infectious ones, have been successfully treated, cancer itself has remained a significant killer, second only to cardiovascular disorders in the West. Cardiovascular death, whether by heart attack or stroke, is generally considered to be part of aging, as it is due to a widespread deterioration in the circulatory system. If it does occur in middle age, it is attributed to excessive stress, over-eating or lack of exercise, originating from the victim's lifestyle. Cancer, on the other hand, originates at one locus and spreads to otherwise healthy tissues in the body. It can occur at any age, although it is mainly associated with old age because of the long timescale of 30 to 40 years often required for the disease to exhibit noticeable symptoms and kill the victim. A person who develops cancer is under great psychological pressure, as are his family [1], because of the popular notions of cancer and its effects - for instance pain, immobility, loss of livelihood and death. The fear of the onset of cancer is strong enough to motivate the general public to donate large sums of money to fund cancer research. This in itself is a strong stimulus in solving the riddle of this unpleasant disease, but the economic loss to the community, resulting from the reduced work-span of many cancer victims is also important.

There have been many advances made in the treatment of cancer, resulting in greater life expectancy and better cure rates, but there will never be a complete and effective treatment - or group of treatments - unless the actual mechanism is understood at a biological level. There is a wide range of scientific detail which must be understood to gain this insight: this varies in detail from the gross statistical clues derived from epidemiological research right down to

the electronic properties of the molecules in the body and their interactions with other agents. This thesis deals with some of these submicroscopic properties relevant to the study of cancer: it studies an enzyme, called glyoxalase I, and its inhibitors. Some of these inhibitors have shown promise as carcinostatic agents in their own right, but they may also be used to prevent the enzyme from destroying other cytotoxic agents used in chemotherapy.

This chapter presents a brief overview of what cancer is, how it can be prevented and treated, and the rôle of the chemist, and especially the theoretical chemist, in this investigation. The second chapter discusses enzyme inhibitors with relevance to cancer research and cancer treatment and the third is a continuation of this theme with reference to the enzyme, glyoxalase I, studied in this work. Chapter 4 contains a discussion of the theoretical methods used in this work and Chapter 5 has details of the computer graphics work which was undertaken to interpret and complement the theoretical results. Chapter 6 presents the information obtained on the catalytic mechanism of the enzyme, based on a set of model reactions and this information is compared with calculations on the inhibitor molecules in Chapter 7.

1.1 The Stages Of Carcinogenesis

The development of neoplastic tumours is a complicated, multi-stage process which is thought to be divided into three main phases [2]: initiation, promotion and progression. This three-stage, multi-step concept of carcinogenesis is based on data obtained from animal tests, particularly on mouse skin tumours, and there is still some debate about its relevance to human cancer, which is not so well

investigated for obvious ethical reasons. The first step is almost certainly a change in the genetic information present in the cell and produces a cell which is indistinguishable from a normal cell, but has the capability of transforming into a neoplastic cell. This conversion will not take place unless there is a subsequent application of the promoting agent [3,4]. Whilst a single application of an initiator is enough to initiate the tumour cell, it seems that repeated contacts with the promoting agent are necessary in order to achieve the conversion to the prolific cell division associated with the neoplastic state. The third stage, that of progression, is the evolving of the benign tumour into a malignant form which can spread - metastasise - throughout the body, forming secondary tumours which ultimately kill the victim.

Initiation is most effective at certain times in the cell cycle: generally the best time to effect the change is at the start of DNA synthesis. This change is directly dependent on the dose of the initiating agent. In contrast, promoting agents have a threshold dose, giving a maximum response. Promotion is a reversible process and is not additive in its effect; in other words, if the application of promoter is dilute enough and spread out over a long enough period, no neoplastic tissue is generated. Thus, the promoting agent can be defined as one which alters the actual expression of the genetic material in a cell. It does not interact with the genetic message itself, as the initiator does, but with the associated decoding machinery. As this stage is reversible, it is an obvious target for cancer treatment: any treatment which can prevent the alteration of the genetic expression, or restore the original coding information will combat the promotion of tumours.

1.2 The Treatment And Prevention Of Cancer

There are many cancers that can be prevented or cured and treatment is improving all the time. Percentage survival rates for many cancers have increased, certainly none have decreased! Certain cancers give rise to early physiological events suitable for screening: the two most obvious examples are cancers of the breast and cervix in women. Cancer of the cervix is almost 100% curable if it is caught early enough. This was not the case twenty years ago. In fact, although the percentage rates of cancer death have increased in the Western World, the absolute values have decreased for any given age group, it is just that the rates of other diseases causing death have decreased even further [5].

As with all diseases, there are two ways to confound the process of cancer: prevention and cure. Judging by current epidemiological evidence, there is great scope for the prevention of many cancers, as the incidence of cancers is not constant throughout the world, nor even related to race; it seems that the pattern of cancer occurrence is defined by the environment experienced by an individual: in other words by occupation, eating, drinking and social habits, social pressures and even the air breathed and chemicals which the individual comes into contact with [6]. That cancer incidence is determined by extrinsic factors to a large extent is best shown by this incidence in a few generations of migrant people. These include Indians who went to live in South Africa, losing their high rate of oral cancer development, Japanese in Hawaii, losing their high rate of stomach cancer incidence and Britons who settled in Fiji, gaining a high rate of skin cancer [5]. Not only do migrant populations lose the pattern of cancer incidence of their native land, they also gain a pattern which is close

to the natives of their new homeland. This is shown by the death rates of Japanese people emigrating to California: in a study comparing death rates of 4 cancers with those of Californian white people it was found that immigrants had death rates more similar to those in Japan than to the Californians for the cancers studied - stomach, liver, colon and prostate cancers - but their sons approach the levels of the whites for these cancers [7]. As the immigrants and children tended to marry within their own racial group, it is fair to argue that it is the changed environment which has affected them. Also, because the death rates of the sons of the immigrants are not identical to those of the indigenous population, it is obvious that some factors such as diet and social habits, which persist into this generation must play a part in the causation of cancer too.

As groups of different culture or different location in the world suffer from different ratios of cancers, it is fair to say that epidemiological research will provide guidelines for living which will help prevent cancer occurrence in a large number of cases. Indeed, some guidelines have already been published [8].

Unfortunately, due to the vagaries and perverseness of human nature, a statement advising against certain practises - even by the government of a country - is not enough to prevent everybody from indulging in those practises known or suspected to give rise to cancer. Therefore preventative measures of this kind are just not enough. There must also be some curative procedure also.

The ideal treatment for any type of cancer would be an agent which selectively removes only the neoplastic cells, leaving the normal ones unaffected. Unfortunately, such an entity does not exist at present. This is because the selectivity of such a drug is not achievable based

on the current knowledge of the structure of neoplastic cells: electron microscopy has revealed that there is generally little difference between healthy and neoplastic cells, especially in the cellular structures common to all cell types [9]. Structures which do change, such as the nucleus, which is often much larger and has an uneven outline, do not consistently exhibit these alterations; indeed, some tumours can consist of cells with normally structured nuclei and others with altered nuclei and yet these have originated from a single original parent cell, assuming that the theory of monoclonality is correct [10]. In principle, surgery can be used to remove the transformed cells, but the problem with this is that there are very few cancers which are detectable early enough to guarantee that the surgeon has eradicated every cell before metastasis can occur.

Thus, the implementation of an effective, reliable systemic chemical treatment is not possible at the moment. Nevertheless, there are chemotherapeutic agents in use in the treatment of cancer. These are universally cytotoxic, as they interfere with DNA, RNA or protein synthesis. They tend to have their greatest success against fast-growing neoplastic tissues such as skin cancers and certain leukaemias; in fact the converse is also true: slow-growing tumours, such as some pulmonary and gastric cancers do not respond very well to these drugs [11]. This is basically because a high turnover of cell division gives the drug a better chance of killing a higher proportion of cells. This is an important point in cancer treatment: to be effective a treatment must dispense with all the neoplastic cells, since even one cell can continue multiplying to give a new neoplasm. Thus, the dose of a chemotherapeutic agent used must be carefully balanced to try to kill all the neoplastic tissue without causing too great a side-effect problem from the depletion of normal, healthy tissues.

There are three main ways of administering chemotherapy to a patient in use nowadays [12]. These are adjuvant chemotherapy, in which the drug is used in combination with some other technique, such as surgery or radiotherapy, to mop up the remaining neoplastic cells, intermittent, high-dose chemotherapy, in which drugs are given in high dosages for short periods, separated by recovery periods, and combination chemotherapy, in which cocktails of more than one drug are used simultaneously in the hope that the effect will be synergistic. The older technique of continuous, low-dose chemotherapy has been almost discarded in favour of the intermittent method.

Perhaps the best method, if at all possible, is adjuvant chemotherapy [13], as it has been shown that chemotherapy's success bears an inverse relation to the number of viable neoplastic cells in the host [10,11]. Therefore, the removal of large quantities of the malignant tissues by surgery or radiotherapy will allow the chemotherapeutic agents a better chance of dealing with the rest.

The situation of cancer treatment nowadays, whilst better than ever before, could not be termed satisfactory. The drugs being utilised at the moment have been discovered by serendipity, or by large derivative screening tests - in other words, by purely empirical means. Few have been designed to attack a particular metabolic process or characteristic of neoplastic tissue. Yet this is obviously the way forward: a careful consideration of the problem at a chemical level will eventually lead to an efficacious treatment. The basic biochemical mechanisms involved both in normal cell division and in abnormal neoplastic cell division must be understood. This is ultimately a genetic problem, as successive generations of tumour cells retain a proliferative ability, but the genetic mechanism is expressed

directly through the creation of polypeptide structures: proteins. Although proteins have a great many functions in the body, ranging from structural, such as membrane proteins, to control, such as hormonal and receptor proteins, their main function with relevance to cancer research is their catalytic role in metabolism enzymes. A proper understanding of the operation of enzymes in neoplastic cell division could lead to a treatment at the root level of the problem. This could be further refined to an examination of active site-substrate interactions and ultimately down to the electronic level. Thus, an understanding of biological mechanisms at the electronic level, as first proposed by Szent-Gyorgyi [14], is needed.

CHAPTER 2

ENZYME INHIBITORS

2.1 Enzymes

Enzymes are the proteins used by all living organisms to catalyse metabolic reactions. The action of these catalysts is specific to single substrates, or to very closely related groups of molecules. The protein part of the enzyme is often inactive unless some other substance, such as a metal atom, is present in the structure. If this co-factor is organic it is known as a co-enzyme, unless it is difficult to separate from the protein, without causing damage, in which case it is a prosthetic group. The inactive protein part is called an apoenzyme and the complete unit is the holoenzyme.

The actual working part of an enzyme is the active site. This is responsible for the chemical manipulation which characterises the enzyme's action; to facilitate this process, there are binding sites to which the substrate is attached. These will arrange the substrate in a certain way in the active site, so that the relevant part of it can interact with the catalytic sites. The active site is only a small part of the enzyme, usually sited on or very near to the surface, so that access is available to possible substrates. The co-factors and amino acid side chains provide strong hydrophobic or hydrophilic areas - microenvironments - which have totally different electrostatic properties from the rest of the enzyme surface, or indeed the surrounding cellular structures. These microenvironments obviously allow the required reaction to proceed easily, by influencing the electrostatic properties of the substrates, by helping the binding of the substrates to the enzyme and by stabilising or de-stabilising various chemical bonds which are being formed or broken.

At some stage in the overall process the structure of the substrate and that of the active site are complementary. Pauling's hypothesis [15] is that this occurs at the transition state between the reactants and products, thus stabilising the formation of this, the most unfavourable step, energetically, in the reaction. The binding of the original substance to the active site must be fairly strong, otherwise the enzyme will never interact with anything, whereas the final binding of product and enzyme will be weak, non-existent, or even repulsive as the enzyme returns to a receptive mode for substrate.

2.2 Enzyme Inhibitors

Inhibitors of enzymes are substances which decrease the rate of the enzyme-controlled reaction. There are two main types: the reversible and the irreversible inhibitors. The first group can be removed by dialysis, or sometimes by dilution, but the irreversible ones cannot.

Irreversible inhibitors bind irreversibly to the active site, often by covalent bonding. This produces an inactivated enzyme-inhibitor complex, which can no longer function properly. The inhibitor may prevent the substrate from binding to the site itself, or it may destroy part of the active site, thus inactivating the enzyme.

Reversible inhibitors are divided further into different categories. These are competitive, uncompetitive and non-competitive. They are distinguished by their mode of inhibition: all of them bind reversibly to the enzyme, forming an equilibrium system of some kind,

which can be studied kinetically, whereas the irreversible inhibitors can increase in inhibitory power with time.

A competitive inhibitor either binds to the same site as the substrate, or binds to a different site, but alters the enzyme in some way - generally in conformation - so that the substrate cannot bind to the active site. The inhibitor-enzyme bonding must be severed before the enzyme can act on its substrate. An example of this is the binding of indole to chymotrypsin, which inhibits the hydrolysis of tryptophan derivatives [16]. The effect of the inhibitor will depend on its concentration and that of the substrate, as well as the affinities of them for the active site. The Lineweaver-Burk plot for competitive inhibition has all the lines converging at the ordinate intercept at 1/Vmax. The Lineweaver-Burk plot [17] is a convenient way of examining the type of inhibition affecting an enzyme. It is a plot of 1/v against 1/s, where v is the initial velocity of the reaction and s is the substrate concentration. This is a straight line with an intercept on the ordinate of $1/V_{max}$ where V_{max} is the maximum velocity achieved when the enzyme is saturated with substrate and an intercept on the abscissa of $-1/K_m$ where K_m is the Michaelis-Menten constant [18].

An uncompetitive inhibitor will only bind to the enzyme-substrate complex. It thus does not compete with the initial substrate in any way, but stops the products from being formed. This is quite uncommon in single-substrate enzyme kinetics - an example is L-phenylalanine inhibiting alkaline phosphatase [19] - but in two-substrate enzyme kinetics uncompetitive inhibition behaviour can be seen if there is no reversible link between the inhibitor and the variable substrate. That is, if the inhibitor is active on the enzyme after both A and B have bound to the enzyme (in that order), then the inhibitor will exhibit uncompetitive inhibition to A. The Lineweaver-Burk plot for this type

of inhibition has parallel lines, with larger values of the ordinate intercept for larger inhibitor concentrations.

The last category is that of the non-competitive inhibitor. This produces an inactive complex irrespective of whether or not the substrate is bound. Therefore, it cannot be binding to the same site as the substrate, which is still free to join to the active centre. The enzyme is unable to process the substrate when the inhibitor is bound, though. The Lineweaver-Burk plot for this has a convergent intercept on the abscissa at $-1/K_m$, as K_m is unchanged by such an inhibitor. Chymotrypsin, which contains a proton acceptor in its active site can be inhibited by decreasing the pH [20]; this is an example of the proton acting as a non-competitive inhibitor, although large changes of pH will change enzyme activity for other reasons, such as the disruption of tertiary structure by ionisation of amino acid side-chains. Large diversions of pH from the enzyme's usual environmental pH can denature it totally and irreversibly.

Another type of enzyme inhibition is allosteric inhibition. This is very common in ordinary metabolism. It is the basis of feedback mechanisms: an efficient utilisation of resources and energy is achieved if the product of a metabolic pathway is also an inhibitor of one of the initial enzymes in the system. This prevents the build-up of unnecessary intermediates and products. It would be difficult for a product, which can be very different from the starting materials, to competitively inhibit the first enzyme in a pathway by binding to the active site. This problem is resolved by having a second binding site on the enzyme with which the substance can interact. The resultant enzyme-inhibitor complex occurs with a change in conformation, disabling the active site properties.

The standard way to state an inhibitor's effect is as an I_{50} value. This is the concentration of the inhibitor needed to inhibit half of the enzyme activity; obviously, the smaller the concentration, the more powerful the inhibitor is.

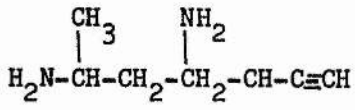
2.3 Enzyme Inhibitors and Cancer

The growth and spread of a tumour is inextricably linked to the enzyme systems used in normal cellular growth. Thus, enzyme inhibitors can have important roles to play in the control of cancerous growth. Some carcinostatic agents have been found to inhibit certain enzymes, and various enzyme inhibitors are known to hinder tumour development.

An example of antitumour agents which are enzyme inhibitors are a series of molecules related to dopamine, reported by FitzGerald and Wick [21]. These are dihydroxy derivatives of dopamine and levodopa and are active against ribonucleotide reductase. This enzyme is involved in the synthesis of deoxynucleoside triphosphates, which are needed for DNA synthesis and is therefore a good target enzyme for control of cell division.

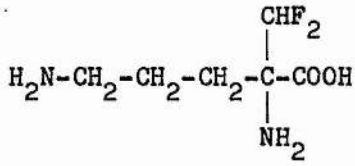
Another target enzyme, which has engendered a great deal of interest, is ornithine decarboxylase (ODC), which is the rate-limiting enzyme in the polyamine biosynthetic pathway. There seems to be a high level of ODC activity induced during the promotion of tumour growth in mouse skin, and the activity correlates with the tumour-promoting activity of various phorbol esters [22,23]. An irreversible inhibitor of this enzyme has been reported: it is (2R,5R)-6-heptyne-2,5-diamine

hydrochloride [24,25].



6-heptyne-2,5-diamine

This was found to prolong the survival of mice with certain leukaemias and lung carcinoma, especially if used in combination with other anti-tumour agents. It was ten times as potent as the previously reported inhibitor, DL-α-difluoromethylornithine [26].

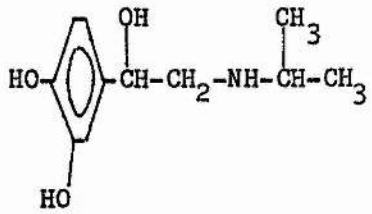


difluoromethylornithine

Development of protease inhibitors was stimulated by the announcement that a protease appeared in mouse skin when 12-O-tetradecanoylphorbol-13-acetate (TPA) and croton oil were applied to it [27]. Proteases are enzymes capable of breaking peptide bonds [28]. They are responsible for digestion of proteins in the gut and for the degradation of proteins in the body, to balance the synthesis of new proteins, maintaining the required concentrations. Protease inhibitors were found to be good inhibitors of croton oil and TPA promotion [27,29]. Also, the inhibitor leupeptin was found to inhibit other colon, oesophageal and mammary tumours in rats [30]. These types of inhibitors can suppress malignant transformation in vitro [31,32],

particularly the promotional stage of the cancer process. Recently, some of these inhibitors have been found to inhibit poly(ADP-ribose)polymerase, too [33]. This is relevant to the study of carcinogenesis because inhibitors of this enzyme can restrict the malignant transformation of cells in both rodents and humans. Poly(ADP-ribose)polymerase is the enzyme responsible for making the polymer poly(ADP)ribose. The role of this substance in the cell is not clear, but it has been associated with regulation of DNA synthesis, RNA synthesis, cell transformation and cell repair mechanisms [34].

Inhibitors of prostaglandin synthesis affect tumour promotion, the best examples being antiinflammatory steroids and dibromoacetophenone, both of which act on the phospholipidase A_2 enzyme, reducing the release of arachidonic acid [35]. This enzyme is near the start of the arachidonic acid pathway, so the production of all end products is suppressed. This is obviously more efficient then stopping the pathway further on, but more subtle control can maybe be gained by inhibiting selectively the lipoxygenase and cyclooxygenase branches of this metabolic route. For instance, some cyclooxygenase inhibitors have been shown to both promote and restrict carcinogenesis, depending on the dose used [36]. Fischer, Mills and Slaga [37] investigated the effects of inhibitors of both lipoxygenase and cyclooxygenase pathways and concluded that control of promotion could be achieved by blocking either the phospholipidase A_2 stage of the pathway, as mentioned above, or both the lipoxygenase and cyclooxygenase ones with appropriate inhibitors. The induction of ODC by TPA is also linked to lipoxygenase and cyclooxygenase: inhibitors of these enzymes hamper this induction in mouse skin, suggesting that some products of arachidonic acid metabolism are essential for tumour promotion by TPA. However, the pathway is not important for ODC induction by isoproterenol, which is believed to involve a cyclic AMP-mediated route [38].



isoproterenol

Another well-studied aspect of cancer research relevant to enzyme inhibitors is that of the mutagenic and carcinogenic metabolites of polycyclic hydrocarbons [39]. The enzyme system which plays a major role in metabolising them is the mixed-function oxidase system, which converts the hydrocarbons into epoxy- and hydroxy-derivatives. These derivatives may be metabolised further by epoxide hydrolase to give hydrodiols, or by UDP-glucuronyltransferases to give glucuronides in combination with glucuronic acid. Some can be acted upon by cytosolic enzymes, such as sulphotransferases and glutathione-S-transferases. The general metabolic fate of these hydrocarbons is to be converted into anionic species, which are more water-soluble and more easily excreted, but activated metabolites lead to the alkylation of genomic DNA.

The polycyclic hydrocarbons are very sensitive to changes in metabolism: For instance, it has been found that 7,8-benzoflavone enhances tumour initiation by dibenz[a,h]anthracene in mouse skin, but inhibits tumour formation by 7,12-dimethylbenz[a]anthracene in the same medium [40,41]. Thus, inhibitors of various enzymes on the metabolic route can shed light on the mechanisms involved and can offer the potential of subtle control of this type of carcinogenesis.

Ireland et al. have studied the repair process which removes benzo[a]pyrene adducts from DNA [42]. They used several inhibitors of DNA topoisomerase, DNA polymerase, ribonucleotide reductase and poly(ADP-ribose)synthetase, which are all thought to play a part in the excision of UV-induced lesions in DNA. It was concluded that benzo[a]pyrene-DNA adducts are removed by a repair process involving topoisomerase, as inhibition of this enzyme led to a much lower rate of removal.

The particular enzyme which we are interested in is glyoxalase I. This has been linked with cancer research since Szent-Gyorgyi proposed it as a promoter of cell growth, in antagonism with its substrate, methylglyoxal, which was the retarding agent [14]. If this enzyme does help to maintain the unrestrained cell division which is a characteristic of neoplastic tissues, then it would be of some use to be able to inhibit it. This is one of the reasons why inhibitors of the enzyme are being sought. Another reason is that some of the inhibitors have been found to restrict the growth of certain cell lines *in vitro*, so it is hoped that more powerful inhibitors of this enzyme will have greater carcinostatic properties.

CHAPTER 3

GLYOXALASE I

This chapter is a continuation of the theme of the last chapter, but is about one particular enzyme, glyoxalase I. It deals with its structure, properties and mechanism and also with the inhibitors of this enzyme. These inhibitors have been implicated as possible target molecules in the design of carcinostatic agents, as some of the first inhibitors discovered were found to have this property.

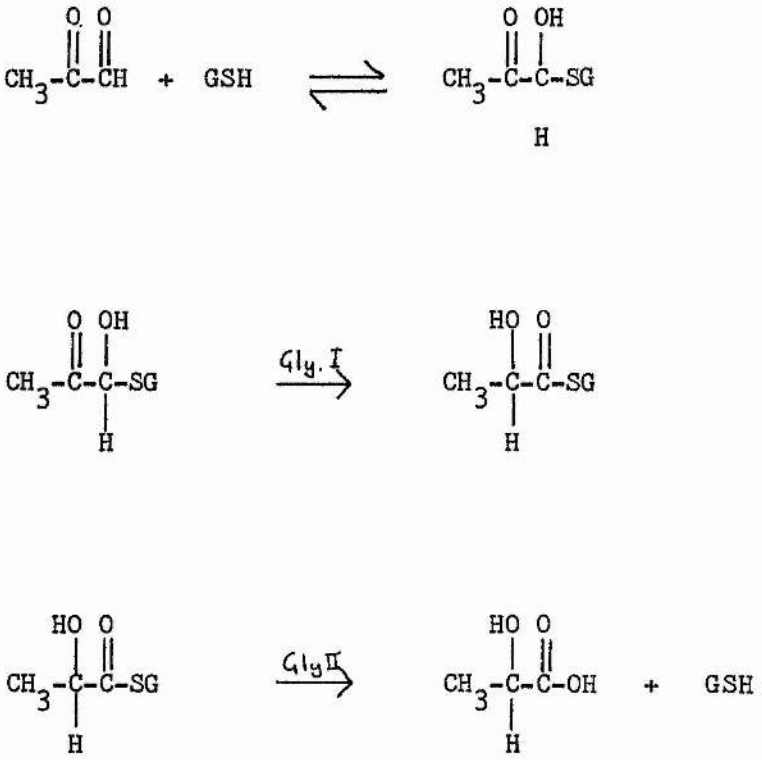
There are two other reasons why this enzyme is linked with cancer research: it has been suggested that the enzyme system may play a role in cell division processes [43,44] or, more simply, that it is a detoxifying system, which destroys certain useful chemotherapeutic agents, which have a structure similar to its substrates, the glyoxals [45].

There have been four review articles written on this enzyme [46-49], the most recent one being an excellent account by Douglas [49].

3.1 The function of glyoxalase

The glyoxalase system consists of two enzymes: glyoxalase I (S-lactoylglutathione lyase, EC 4.4.1.5) and glyoxalase II (hydroxyacylglutathione hydrolase, EC 3.1.2.6). The function of the system is fairly well characterised: glyoxalase I acts on a hemithioacetal, which is formed by the reaction of glutathione and methylglyoxal, or another 2-oxoaldehyde. The enzyme produces S-lactoylglutathione (in the case of methylglyoxal), which is converted

to D-lactic acid by glyoxalase II:



The glyoxalase system was discovered in 1913, independently, by Dakin and Dudley [50] and by Neuberg [51]. It is found in all cell types in the body, and in all species, from microbes to man. However, its physiological significance is far from clear. Various functions have been assigned to it in the past: Gillespie [52,53] has proposed it as an agent in microtubule assembly and anti-(immunoglobin E)-induced histamine release; Aronsson and Mannervik [45] have suggested that it plays a detoxification rôle, protecting the body from intestinal bacteria (though why the bacteria themselves should contain the enzyme is not explained!); others have put forward uses for glyoxalase in the degradation of glycine and threonine [54] and in the heme biosynthetic pathway [55]. Vince and Wadd [43] proposed that glyoxalase plays a

part in controlling cell division, by regulating the amount of methylglyoxal present. At roughly the same time, Szent-Gyorgyi propounded a similar idea, and he has written various articles and books [14,44,56,57] expounding a rather original theory in which methylglyoxal is a growth regulator in antagonism with glutathione, through the medium of glyoxalase. This is linked to cancer and carcinogenesis by his contention that normal cells are in a higher, β state than cancerous cells; these tumour cells exist in the highly proliferative α state, linking them with primitive life forms, which did little else but multiply and survive. These theories have prompted a large body of work on glyoxalase I and its inhibitors, especially as the inhibitors often have carcinostatic properties, though why the inhibitors have this power is not at all clear. The reason that the majority of the work is carried out on the glyoxalase I part of the system is that the products of this first enzyme are generally harmless to the cell: the glyoxalase I reaction is the important step in the detoxification process, proceeding at a rapid pace; the subsequent hydrolysis is carried out in a much more leisurely manner by glyoxalase II.

3.2 Properties of Glyoxalase I

The glyoxalase system occurs in the highest mammals and in organisms as simple as yeast and bacteria. The glyoxalase I enzyme itself exists in different forms: the mammalian enzymes have two subunits, whereas the yeast and bacterium versions have only one. Marmstal et al. have compared the mammalian enzyme to one purified from yeast and have found significant differences [58]. The microbial enzyme has a molecular weight of about 32,000 atomic units and the

mammalian one has a molecular weight of 46,000. It seems that, whilst the microbial glyoxalase I is spherical, the mammalian ones are elongated in shape. This would explain the noted differences in calculated weights for enzymes from various mammalian sources, determined by different physical methods.

The chemical structure of these two enzymes differ, too; apart from the presence of two subunits per enzyme in mammals, there is a great variation in proportions of amino acids, as the isoelectric point of the mammalian glyoxalase I is at a pH of 4.8, whereas that of yeast is at a pH of 7.0. The mammalian enzymes from distinct origins (rat, pig and human) have slightly different numbers of amino acids, but are all very different from the yeast enzyme. It seems that the mammalian enzymes have a common evolutionary origin, but that the microbial one has evolved separately to do the same task. Notwithstanding these differences, the two enzymes have similar k_{cat}/K_m values and substrate specificities and mechanisms. This suggests that the active centre and basic catalytic mechanism are identical - or very similar - for these two types, even though they evolved separately. There is some conflicting evidence on this point, however, as the yeast enzyme is susceptible to thiol-blocking reagents [59], which do not affect the mammalian version. Whether this is due to some variance in the reactive site of the microbial protein, or just to the surrounding structure, which affects the catalytic properties in an allosteric-type interaction, has not been resolved.

The human enzyme can exist in three isoenzymic forms: these have been called α^2, $\alpha\beta$ and β^2, as it has been suggested that they arise from the joining together of two slightly different subunits [45]. This could explain the rather contorted kinetics data which have been reported in the past [59-61]: The active sites on different subunits

could act in a cooperative manner, giving a non-Michaelian rate equation and non-linear inhibition by glutathione.

Glyoxalase I was originally thought to be a Mg^{2+}-dependent enzyme [62], but it became obvious that the metal ion involved was, in fact, Zn^{2+} [63]. The apoenzyme is inactive in both yeast and mammalian glyoxalase I, but unlike the yeast version, the mammalian enzyme can be re-activated by about a ten-fold excess of not only Zn^{2+}, but also Mn^{2+}, Mg^{2+} and Co^{2+} [64]. The reactivated enzymes were found to have essentially the same K_m value for the catalysis of methylglyoxal as the native enzyme. However, the inhibition constants of the other metallo-substituted enzymes increased relative to zinc-glyoxalase I, which had the same inhibitor constant as the native one, showing that zinc is indeed the metal present in glyoxalase I. It was concluded that the zinc must be situated somewhere in the active site, since it was necessary for catalytic activity.

In the same article, Aronsson et al. investigated the presence of tryptophan residues in the enzyme, by means of fluorescence spectroscopy and the tryptophan-modifying reagent, 2-hydroxy-5-nitrobenzyl bromide. They found that there were two tryptophan residues per molecule that could be shielded from this reagent by S-p-bromobenzylglutathione - a good inhibitor of the catalytic process. This implied that there was one tryptophan residue present at the active site of the enzyme, since glyoxalase I consists of two almost identical subunits, each possessing an active site.

The same group has also investigated the effect of sulfhydryl and amino group reagents on glyoxalase I [65,66]. Whilst yeast glyoxalase I is inactivated by various sulfhydryl reagents, including mercaptide-forming, alkylating and oxidising reagents, glyoxalase I

derived from porcine erythrocytes was unaffected, except by mercuric chloride and p-mercuribenzoate. (These reagents do affect the reaction by destroying glutathione, however). Both enzymes reacted in a similar manner to amino group reagents: acetic anhydride and pyridoxal failed to impair the catalysis but dansyl chloride, 1-fluoro-2,4-dinitrobenzene and 2,4,6-trinitrobenzene sulphonate were effective. It was established that these reagents were not acting at the same place as the sulfhydryl ones, by looking at the effect of multiple inactivation: if the two inhibitors are acting at the same site, they would compete for it, and the resulting impedition would be less than the sum of the separate reagent's action. It was found that some of the inactivators, for instance 2,4,6-trinitrobenzene sulphonate and N-phenylmaleimide, acted synergistically. Inactivation of the enzyme, followed by attempted reactivation by addition of excess glutathione showed a further difference between the actions of p-mercuribenzoate, 2,4,6-trinitrobenzene sulphonate and a mixture of the two: the first compound gave almost complete inactivation, followed by almost full reactivation, whereas the second one was unaffected by the glutathione. When both were added, a stronger inhibition resulted, and the glutathione did not reactivate the enzyme effectively. Since the mammalian enzyme is not affected by sulfhydryl reagents to any great extent, it can be concluded that sulfhydryl groups are not directly involved in the catalytic process of the enzyme. Also, they do not react at the same site as the amino group reagents, which seem to react in a similar manner on both enzymes; the effect of increasing the pH is to increase the apparent first-order rate constant, implying that the basic form of a nucleophilic group is modified. The authors rationalised the evidence by suggesting that there were amino groups present at the glyoxalase I active site.

Glyoxalase I exhibits similar catalytic properties when the zinc atoms are replaced by other divalent metal atoms [63]. Therefore, the magnetic properties of other metals can be used to scrutinise the active centre. The manganese atom is paramagnetic, and so is suitable for study by NMR techniques. The relaxation rates of water protons interacting with manganese were studied by Sellin et al. [67,68] using NMR spectroscopy. The manganese-glyoxalase I was found to have about 70% of the catalytic power of the native enzyme and had the same fluorescence spectrum, due to tryptophan residues, as mentioned previously [64]. Conclusive proof that the metal was part of the active site of the enzyme was established; this had been assumed previously, but never proven: the metal might have just been a structural feature of the enzyme. The paramagnetic influence of manganese on protons of glutathione and S-D-lactoylglutathione was observed, showing that these bind very close to the metal. As glutathione is part of the hemithioacetal substrate, and S-D-lactoylglutathione is the product of the reaction, the metal must be part of the active centre.

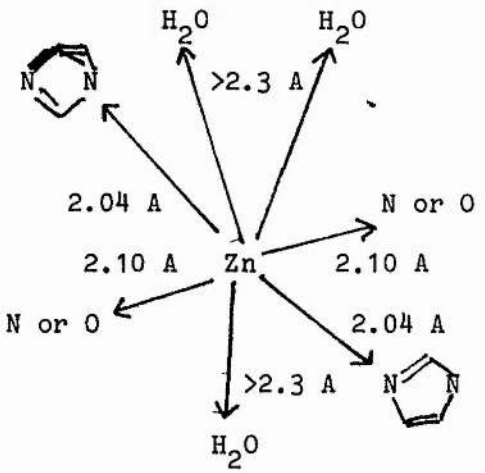
It was calculated that there were two exchangeable water ligands attached to the manganese in the unbound state, assuming that the metal had octahedral coordination. In the presence of the inhibitor, S-(p-bromobenzyl)glutathione, there is only one water ligand observed to be exchanging. This implies that either the water ligand is displaced by the inhibitor, or, as seems more likely, that the water is immobilised by the binding of the inhibitor. This was found to be the case for S-D-lactoylglutathione, too.

From their results Sellin et al. proposed a mechanism for the action of glyoxalase I on a hemithioacetal. This is pictured in Figure 3-1. The reaction mechanism suggested is as follows: the hemithioacetal is bound in the active site, forming a hydrogen bond between its keto group and one of the water ligands of the metal. This favours the protonation of the keto oxygen and the simultaneous abstraction of a proton from the C1 carbon. This results in an enediol with a negative charge on the C1 oxygen, which could be partly neutralised by hydrogen bonding to the second water molecule. The final step is the donation of the proton from the erstwhile abstracting base to yield the product. The enzyme has two distinct conformations: the resting state and the binding state. It changes from the resting state to the binding state upon interaction with the hemithioacetal, or the S-substituted inhibitor: this is confirmed by the observation that the fluorescence spectra of the bound enzyme and the inhibited enzyme differ from the unbound variety. The catalytic action then restores the enzyme to the resting state, and the product dissociates: the S-D-lactoylglutathione fluorescence spectrum is the same as the unbound one.

Another technique, using properties of the metal in the enzyme is that of the X-ray absorption, applied by Garcia-Iniguez et al. [69]. This furnished more information on the coordination of the metal, and on likely ligands thereof. The X-ray absorption spectra of glyoxalase I and of glyoxalase I plus the inhibitor S-(p-bromobenzyl)glutathione were very similar to that of $(Py)_3Zn(NO_3)_2$, a 7-coordinated Zn^{2+} compound. The X-ray spectra of the $Zn(ImH)_4(ClO_4)_2$ and $[Zn(ImH)_6]Cl_2$, which were 4- and 6-coordinated Zn^{2+} models respectively were quite different. This implies that the glyoxalase I zinc has a distorted octahedral geometry. Also, the Zn^{2+}- N or Zn^{2+}- O distances calculated

from the fine structure data are larger than would be found for tetrahedral complexes. No evidence was found for the glutathione sulphur binding to zinc.

Back-scattering from atoms at about 3 to 4 A from the zinc were similar to those found for known zinc-imidazole models, including carbonic anhydride, a copper-zinc metalloenzyme with imidazole ligands. The implication was that at least two imidazoles were bound to the metal. Additional conclusions led to the proposed structure below, with two imidazole ligands, two N- or O-containing ligands and two or three waters:



Recently, some more NMR work, in conjunction with computer modelling techniques, has provided even more detailed evidence on the nature of the binding in the enzyme [70]. Rosevear et al. studied the conformations of four glutathione derivatives bound at the manganese-enzyme active site. These derivatives were a substrate analogue, S-(acetonyl) glutathione, a hydrophobic inhibitor, S-(propyl) glutathione, a charged inhibitor, S-(carboxymethyl) glutathione and the product, S-(D-lactoyl) glutathione. The paramagnetic effect of the Mn^{2+} atom on the nearer protons was used at various frequencies to calculate distances from metal to proton. These were used, along with

^{13}C NMR results for the product, as parameters for a computer fitting program in order to calculate the conformations. Each of the glutathione derivatives adopted a Y-shaped conformation when bound to the enzyme, similar to the theoretically predicted structure [71] of glutathione, itself.

The substrate analogue, S-(acetonyl) glutathione differs from the actual substrate in that it has two protons on C1 instead of one proton and a hydroxy group. The best fit calculated conformation from the NMR data has the carbonyl oxygen pointing towards the metal as one of the constraints, as this had been found to be the case previously [72,73]. The C1-C2 bond is more or less perpendicular to the C-Mn direction and the sulphur atom points towards the Mn. The C1 protons are on the opposite side from the metal, as is the methyl group. The distance from the carbonyl oxygen to the manganese atom was found to be 5.1 A, whereas a manganese-oxygen distance of between 2.0 and 2.4 A would be expected for direct coordination. This supports the earlier work, suggesting an intermediary water ligand and second-sphere coordination of metal to oxygen.

The product conformation exhibited some notable changes from the substrate analogue: the lactoyl carbonyl oxygen pointed towards the metal and the sulphur away from it, in contrast to the substrate analogue; the methyl group had a similar position to before, whereas the hydroxy oxygen, being joined to a tetrahedral carbon now, was not pointing directly at the metal, being at an angle to it. The greatest difference between the two structures was at the C1 location, which was, on average, 1.6 A away from the corresponding atom in the substrate analogue. This difference was the only one that was greater than the calculated uncertainty in the positions of the atoms of both species and implied a conformational change from substrate to product.

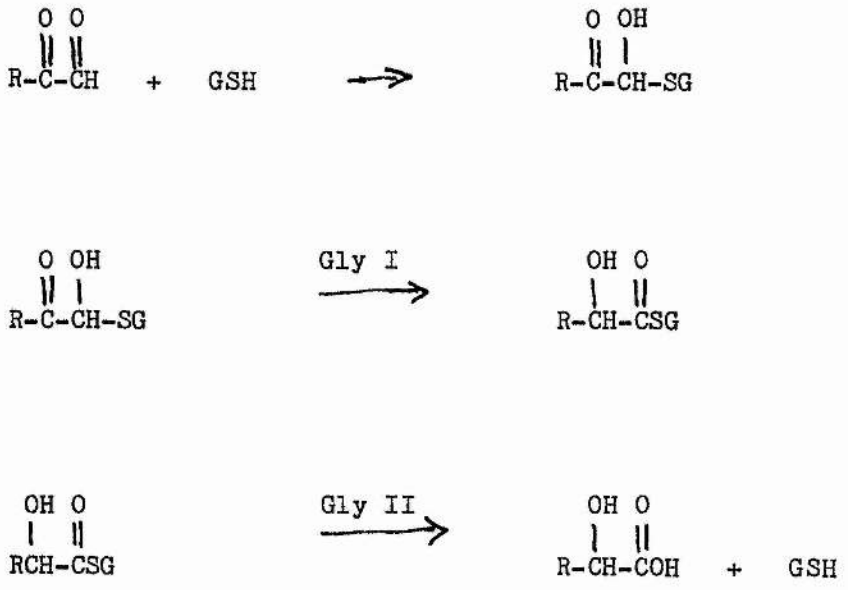
This can be explained either as a flexibility in binding of the enzyme to allow both stereoisomeric substrates to bind, as reported by Griffis et al. [74], or as conformational changes in the enzyme to bind differently with substrate and product [68].

S-(propyl) glutathione is a hydrophobic inhibitor of glyoxalase I - at least, the propyl end of the molecule is hydrophobic. The calculated structures for this compound were more variable than the others, perhaps because there was no sp^2 carbon atom to constrain the structure. The best fit structure is fairly similar to the acetonyl-substituted glutathione, with C1 and C2 roughly the same distance from the metal and the C3 methyl group pointing away from it.

The charged substituent is further away from the metal than the others: if they are considered to be second-sphere coordinated, then the S-(carboxymethyl) glutathione is third-sphere. It still has the extended Y-shape of the glutathione moiety present, as do the other compounds, with no internal hydrogen bonding.

3.3 The catalytic mechanism of glyoxalase I

The reaction scheme below is a simple representation of what the glyoxalase system does:



A hemithioacetal is formed by the α-ketoaldehyde and glutathione. This is converted to an α-hydroxythioester by glyoxalase I, this thioester being further hydrolysed to the corresponding acid by glyoxalase II.

The important part of this process, as far as detoxification is concerned, is the glyoxalase I reaction, because the products of this reaction are non-toxic in all known cases [75]. The reaction catalysed by this enzyme is a hydrogen transfer from the hydroxy-substituted carbon to the keto carbon and a corresponding isomerisation of carbonyl and hydroxy groups. This hydrogen transfer step is the most interesting in the reaction; there have been conflicting reports on the nature of this process over the last thirty years. Both intramolecular hydride transfer and intermolecular proton transfer mechanisms have been suggested in the past [76-78]. The hydride transfer is analogous to a Cannizzaro reaction in strong alkali. This seemed to be supported by the fact that no C-D stretch was observed by infra-red spectroscopy when the reaction was performed in D_2O [76,77], and that almost no ^3H

was incorporated when the reaction was carried out in tritium-enriched water [78]. A model reaction, using hydroxide ion catalysis on phenylglyoxal in aqueous solution has been investigated by both Hine et al. [79] and Vander Jagt et al. [80]. These reactions were proved by both groups to be rate-limiting hydride-transfer reactions, but there were notable differences from the glyoxalase I reaction itself. As Vander Jagt et al. pointed out, the high sensitivity to substituent, which was a feature of the model reaction was in marked contrast to the lack of sensitivity to substituent in the glyoxalase I reaction. The model reaction followed a Hammett relationship with ρ 2.0, which indicates a transition state stabilised by electron-withdrawing groups. This was not true for glyoxalase I.

Another model scheme investigated has been that involving a general base catalysis mechanism, reported by Hall et al. [81,82]. In the presence of magnesium nitrate hexahydrate, or other divalent metal salts, and a tertiary amine or sodium acetate, the rearrangement of α-ketohemimercaptals to the equivalent α-hydroxythioesters was studied in DMF. It was found that the rate of rearrangement was increased about 30-fold over the non-catalysed reaction [81]. This scheme was modified to mimic more accurately the glyoxalase system: glutathione, magnesium ions and either imidazole or phosphate dianion were used. Both imidazole and phosphate were found to be effective, but primary amines and thiols were considered to be too strong, leading to nucleophilic adducts with α-ketoaldehydes.

The rate was increased 1.8 times when Mg^{2+} was added to the imidazole reaction. (In a non-aqueous medium, as stated above, a 30-fold acceleration was noted, but, in this system, water probably competes for Mg^{2+} complexing sites with the imidazole). The kinetic isotope effect was much lower than for the hydroxide ion catalysed

Cannizzaro mechanism. This seems to differentiate the proton transfer enediol mechanism from the 1,2-hydride shift, for the model scheme. However, $V_{max}(^1H)/V_{max}(^2H)$ varies with the source of the enzyme, so all that can be confidently concluded is that the hydrogen transfer is the rate-limiting step, but not what type of transfer it is.

Control tests were also carried out to establish whether the substrates, intermediate or product exchanged protons with the solvent: it was found that they did not, by NMR studies, but the glyoxalase catalysed reaction in D_2O and the model reaction did exhibit solvent hydrogen incorporation. The Cannizzaro reaction in strong base was also shown to be a hydride-transfer mechanism as the lactic acid product from methylglyoxal in D_2O exhibited a ^1H signal. Lastly, another control experiment was carried out, proving that even in strong base, the product did not incorporate deuterium at the carbon.

Okuyama et al [83] investigated the equilibrium and kinetic properties of the reaction between thiols and aldehydes to give hemithioacetals, and the subsequent rearrangement to thioesters, by spectrophotometric means. They found that K_h, the equilibrium constants for the reactions, follow sigmoidal curves, of pKa = 9.0 for glutathione and pKa = 9.7 for 2-mercaptoethanol. These correspond to the pKa of the thiols themselves, implying that just the neutral thiol species reacts to form the hemithioacetal at equilibrium.

The rates of rearrangement were found to be dependent on RSH concentration, with saturation setting in at higher concentrations. These were found by studying the slow decrease in absorbance in the 280 nm region, owing to hemithioacetal being converted into ester. The K_h obtained by this method agreed well with that obtained for the hemithioacetal equilibrium, confirming that this is an intermediate in

the reaction.

It was also established that the rearrangement relied on general base catalysis, as the rate constants k_{max},- corresponding to those of the rearrangement when all the aldehyde is transformed - increase with greater buffer concentrations. This supports the enediol intermediate and proton transfer mechanism. The solvent isotope effects pointed to a rate-determining step in the rearrangement being a proton transfer from the hemithioacetal to the base, too. The rates of rearrangement seem to increase with increasing pKa of the thiol: this is probably because electron-withdrawing groups, which lower the pKa, also facilitate deprotonation of the hemithioacetal.

Flavins can trap transient enediol intermediates oxidatively [84-86], but are not very good hydride acceptors [87]. Shinkai et al. [88], used 3-methyltetra-O-acetylriboflavin to try to trap any enediol intermediate formed in the glyoxalase I model reaction. The flavin almost completely inhibited the formation of the ∝-hydroxythioesters from hemithioacetals. This supported the contention of Hall et al., that glyoxalase I catalyses a proton abstraction/donation mechanism. The oxidation of flavin was zero-order in flavin, but first-order in both thiols and glyoxals; therefore the rate-limiting step is the deprotonation to form the enediol intermediate and the flavin oxidative trapping is relatively rapid.

Of course, the mechanism might be an intramolecular hydride shift to form the product, which <u>then</u> forms an enediol by rearrangement and loss of a proton. This mechanism was excluded by seeing if the product analogue, S-ethyl thiomandelate would react with the system to give any flavin trapped enediol. It was found that the compound was converted to mandelic acid almost totally, with no benzoylformic acid being

detected. Therefore, the hydrolysis of the thioester is much faster than the deprotonation to yield the enediol. Also, when S-ethyl thiomandelate was mixed with the flavin in a buffered solution, no reduction of the flavin was observed. Thus, the presence of the hemithioacetal was a necessary condition for the reduction of flavin. The authors concluded that their results confirmed the earlier NMR work [81,82], that the rearrangement of α-hemithioacetals to α-hydroxythiol esters proceeds via 1,2-enediol intermediates in the analogue of the glyoxalase I catalysed reaction. The authors also attributed stabilisation of the carbanion produced by the proton abstraction as a rôle for the sulphur atom.

More recently, this work has been extended to flavin trapping of an enediolate intermediate in the yeast glyoxalase I reaction, itself [89,90]: the flavins were found to be weak inhibitors of the enzyme; this is an ideal state of affairs for trapping any intermediate, as stronger inhibitors would probably obstruct the reaction mechanism too much. A control experiment without glyoxalase I present showed that the flavins were slowly reduced by the substrate mixture, but the reduction rate in the presence of the enzyme was an order of magnitude greater. The rate, after allowing for the background reduction, was found to be dependent just on the amount of glyoxalase I present in the reaction mixture. The reaction was again zero-order in flavin concentration, as in the model reaction, as the hydrogen abstraction is the rate determining step, followed by a fast reaction with flavin. At least, this is true for low levels of flavin, but at higher levels, there was some dependence of the reduction rate on concentration. This was proposed to be due to partitioning of the flavin between the solvent and the active site, which is probably hydrophobic. Flavins would certainly be present at the active centre in some proportion, because they are weak inhibitors of the enzyme. Further evidence on

this point was gleaned from the fact that the more hydrophobic flavins were reduced at a greater rate and in greater number than the hydrophilic ones, suggesting that the hydrophobic flavins were possibly at a large hydrophobic site near to the substrate. As mentioned before, the flavins do not trap hydride ions very well, so this work advocates a place for a carbanionic enediolate intermediate in the enzyme-catalysed reaction, especially as the apparent K_m from the flavin reaction is very similar to that obtained by Vander Jagt et al. [91] for the normal enzymic reaction, signifying that the enzyme is probably functioning normally in the flavin experiments.

3.4 Inhibitors of glyoxalase I

The substrates of glyoxalase I, such as methylglyoxal, are cytotoxic in appreciable amounts, but may be regulatory in small concentrations. This is the basis of the suggested use of glyoxalase I inhibitors as anticancer agents: if glyoxalase I activity is reduced, there may be a build-up of its substrates, thus providing a means of regulation of cell division in neoplastic cells.

3.4.1 Glutathione based inhibitors

The natural substrate for this enzyme consists of a hemithioacetal formed from the glyoxal and glutathione, so various derivatives of glutathione were designed and synthesised to probe the mechanism of the enzyme. These were the first inhibitors to be designed based on the mechanism of the catalytic reaction of glyoxalase I and may be thought

of as substrate analogues.

S-substituted glutathiones have been prepared by various groups [92-95]. They are inhibitors of glyoxalase I, and many were found to kill L1210 leukaemia and KB cells [96].

Lyon and Vince [97] have investigated a range of S- and N-substituted cysteinylglycines for antagonistic activity towards yeast glyoxalase I. This work provided interesting new data on the binding of substrates to glyoxalase I: contrary to expectation, these glutathione analogues did not show competitive inhibition properties, as had S-substituted glutathiones previously, but were non-competitive inhibitors of yeast glyoxalase.

When the S-benzyl group was replaced by S-p-bromobenzyl, binding was found to be much stronger. Additionally, S-(p-bromobenzyl)-L-cysteinylglycine was found to be inhibitory, whereas S-benzyl-L-cysteinylglycine was inactive. These facts point to a hydrophobic site accessible to the S-substituent of glutathione.

Kinetic experiments were run to ascertain the nature of the noncompetitive binding of the cysteinylglycines. These established that the hemithioacetal, glutathione (which is a weak competitive inhibitor itself), and S-(p-bromobenzyl)glutathione all bind at the same site. The inhibitor, glutaryl-S-(p-bromobenzyl)-L-cysteinylglycine, did not compete with glutathione for this site, binding elsewhere, with a slightly synergistic effect. So, a small change in the glutamyl part of glutathione allowed good binding, but prompted the molecule to bind to a different site.

3.4.2 Natural inhibitors

Oray and Norton [98] investigated the inhibition of both glyoxalase I and II by nucleotides and nucleosides. They found that inhibition was constant for a nucleoside, its mono-, di- and triphosphates and for the free base. Their results suggest that the glyoxalases would be completely inactive *in vivo*, because of the concentration of nucleotides present in cells. Some sort of activator would be required to reduce this inhibition, or the enzyme would have to be segregated in a vesicle, away from the general cell contents.

A strong inhibitor of glyoxalase I has been separated from a culture of Streptomyces griseosporeus [99]. This was found to be 2-crotonyloxymethyl-4,5,6-trihydroxycyclohex-2-enone. Another culture, this time from a mushroom, Stereum hirsutum, provided a larger inhibitor molecule called MS-3 (3'4'-dihydroxymethyl-5'-hydroxy-6'-(3-methyl-2-butenyl)-phenyl-2,4-dihydroxy-6-methyl benzoate) [100], whose structure was elucidated by X-ray crystallography, as was the previous molecule. These compounds are interesting enough in their own right, but a more systematic approach as detailed below gives more information on the nature of the catalytic process.

3.4.3 Reductones

Iio et al. [101], studied the effect of a series of chemical species, which they termed reductones, on glyoxalase I, as these have antitumour properties. These chemicals are usually known as antioxidants as they are involved in the control of reactive oxygen

species in the body. Iio and his co-workers found that ascorbic acid and ascorbic acid 3-phosphate both inhibited glyoxalase I well, but dehydroascorbic acid was not inhibitory. Aromatic reductones all inhibited the enzyme, but were about ten times weaker than the aliphatic ones. Reductones have an enediol structure, which was claimed to be essential for anticancer activity, but not for inhibition of glyoxalase I, as ascorbic acid 3-phosphate was an inhibitor, even though the enediol structure was not present. Ascorbic acid was found to exhibit uncompetitive inhibition, that is, it interferes with the enzyme-substrate complex.

The group found that p-benzoquinone was a strong inhibitor of this enzyme. This was rationalised as being due to four simultaneous processes taking place; these were: attack of the inhibitor at enzyme sulfhydryl groups or amino groups, reaction with glutathione, competitive inhibition by oxidised glutathione and competitive inhibition by glutathione-benzoquinone adducts.

The authors concluded that part of the reductones' anti-cancer effect may be due to their inhibition of the glyoxalase system, in addition to their interference with DNA chains.

3.4.4 Mechanism-based inhibitors

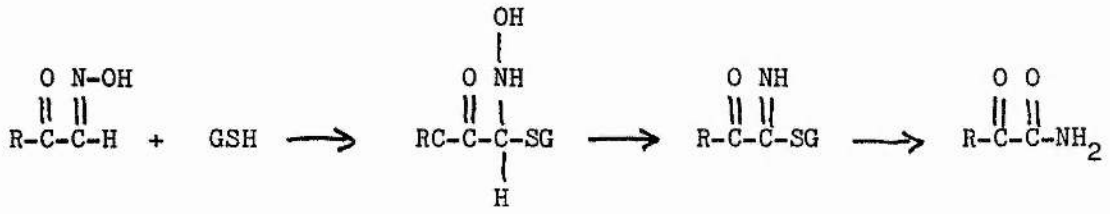
Douglas and Nadvi [102] also placed emphasis on the enediol structure of inhibitors not related to the substituted glutathione compounds. They based their approach to the problem of inhibitor design on the mechanistic studies that showed that an enediol intermediate was present in the reaction. Maltol was found to be a

non-competitive inhibitor and was as good an inhibitor as the substituted glutathiones. Like the ascorbic acid 3-phosphate mentioned above, this does not contain an enediol structure, but the authors pointed out that all the inhibitors contained either an enediol or a paene-enediol structure, that is a structure similar to that of the enediol. Whether the inhibition was due to the similarity of inhibitor to the Michaelis complex (i.e. substrate- enzyme complex) or to the enediol transitory species could not be ascertained. The inhibition kinetics of the other inhibitors, such as squaric acid and 2,3-dihydroxybenzoic acid were not reported. The reason for the observed non-competitive inhibition by maltol may be that the inhibitor is interacting with part of the active site, allowing the substrate hemiacetal to bind normally, but blocking part of the catalytic mechanism. Thus, if the glutathione part of the substrate, - or possibly just the cysteinylglycine part - is responsible for attaching to the enzyme, bringing the hemiacetal part close to the catalytic centre, then obstruction of the catalytic part of the site by maltol would not affect the binding, just the subsequent reaction.

This paper prompted a great deal of study on enediol and paene-enediol compounds as possible glyoxalase inhibitors, leading to the discovery of some very potent inhibitors, such as the flavone series mentioned below. Various groups have since tried to find inhibitors of glyoxalase I by mechanistic reasoning, based on the structure of parts of the putative inhibitors.

In a review article, Jordan et al. reported the inhibition of the enzyme by methyl- and phenylglyoxaldoxime [48]. These compounds were investigated for inhibitory activity because they are good chelators and are similar in structure to α-ketoaldehydes. They turned out to be fair inhibitors, with 50% inhibition at a concentration of about

0.004 M. In the presence of thiols, the oximes rearranged to α-ketoamides. Glyoxalase I was found to catalyse this reaction and the following reaction scheme was proposed in explanation:



The addition of oxime to glutathione is fairly unfavourable, being at least 300 times slower than for the corresponding α-ketoaldehyde. This led the authors to suggest a two-substrate mechanism in which addition takes place after binding to the enzyme. The next step is the deprotonation by an enzymic base, analogous to the catalysed rearrangement of the aldehydes. Finally the adduct was hydrated, to remove the glutathione and form the amide.

This investigation provided supporting information for the enediol mechanism, but the oximes were not very good inhibitors. This was construed to be due to two hindrances on the reaction: the addition to glutathione is not favourable and the oximes are probably found in a trans conformation, which does not favour interaction with the enzyme active centre.

Douglas and Ghotb-Sharif [103] reported the inhibition of yeast glyoxalase I by a series of tetrapyrroles or porphyrins. They found that the enzyme was inhibited by these, but not by vitamin B-12. Despite the fact that haemin was found to react with glutathione, kinetic measurements showed that the inhibition was probably by direct

action on the enzyme, not by glutathione depletion. The inhibition was also shown to be of a competitive nature. The binding to the enzyme was found to be pH-dependent, too. The pH profiles gave a pK_{app} of around 7, which was probably an enzymic ionisation, as the pK_a values for the side-chains of the porphyrins were much more varied.

It seems that the enzyme must have a large hydrophobic pocket close to the active site to accommodate these big molecules, as porphyrins have a hydrophobic centre, surrounded by ionic side-chains, which presumably interact with the active centre in some way, hindering the enzymic catalytic process.

Brandt et al. [104,105], investigated some of the inhibitors discussed above, along with some new ones, using a glyoxalase I prepared from human erythrocytes - probably the most relevant system hitherto used for the study of the carcinostatic and metabolic properties of the enzyme and its inhibitors. The new compounds were derivatives of coumarin, that is esculin, esculetin and 4-methylesculetin. Esculetin and 4-methylesculetin had the same I_{50} value of 0.03 mM and esculin was a weaker inhibitor, at 0.23 mM. When esculetin, 4-methylesculetin and squaric acid were compared, by plotting graphs of percentage activity of the enzyme against the concentration of the inhibitor, it was found that the inhibition curves were different, suggesting that the type of inhibition may vary between these compounds.

A number of flavone derivatives were found to be very powerful inhibitors of glyoxalase I by Brandt et al. [106]. These were hydroxy-substituted derivatives of 4',5,7-trihydroxyflavone - apigenin. The best inhibitor was myricetin, for both human erythrocyte and yeast glyoxalase I. This has a trihydroxy-substituted phenyl

ring. When myricetin is modified at the 3-hydroxy position to give its glucoside, myricitrin, the inhibition dropped dramatically, from an I_{50} of 5.0 uM to 185.0 uM. This is true for the other flavones, also. The extent of hydroxy-substitution on the phenyl ring also seems to be important, as this is the only difference between myricetin and quercetin, morin, kaempferol and 3-hydroxyflavone, which show a range of I_{50} values of one order of magnitude. There is an interesting point arising from the inhibitory power of these compounds: 3-hydroxyflavone seems to be inordinately good as an inhibitor; the rest of the compounds decrease in activity with decreasing hydroxy-substitution on the phenyl ring, yet 3-hydroxyflavone, which has no hydroxy-substitution, is as good as quercetin and fisetin.

A related group of molecules - substituted coumarins - also exhibited inhibitory prowess [107]. These again were hydroxy-substituted compounds. Coumarin itself is not inhibitory to any great extent, but some of the derivatives are very good. The best inhibitors have two hydroxy groups adjacent to each other, reminiscent of the cis-enediol structure discussed above. A non-polar group at the 4-position also seems to enhance the inhibitory effect of these compounds. It was also found that compounds which had no enediol strucure, or had a keto group next to the hydroxy one were only weakly inhibitory. The type of inhibition was found to be competitive.

The above groups of inhibitors, namely the flavones and the coumarins, are ideal for study by theoretical means: they are two groups of similar molecules, with a range of inhibition of over one order of magnitude and contain the simple elements carbon, hydrogen and oxygen, which allow the use of minimal basis set calculations to give fairly reliable results. An ideal method of studying the electronic properties of these compounds is to look at these properties on their

van der Waals surface, as the similarity in shape and size allows detailed comparison of the structures. The methods of study will be treated in depth in the next chapter.

This field is an ideal one for the interplay of theoretical and practical disciplines, as the stimulus of the experimental work has led to a careful theoretical consideration of the mechanism of this enzyme. This has in turn provided the experimentalist with new compounds to synthesise and test, based on the calculated properties of proposed inhibitors. The cooperation of another NFCR research scientist, Professor Richard Brandt, in the testing of molecules has been an important factor in this project.

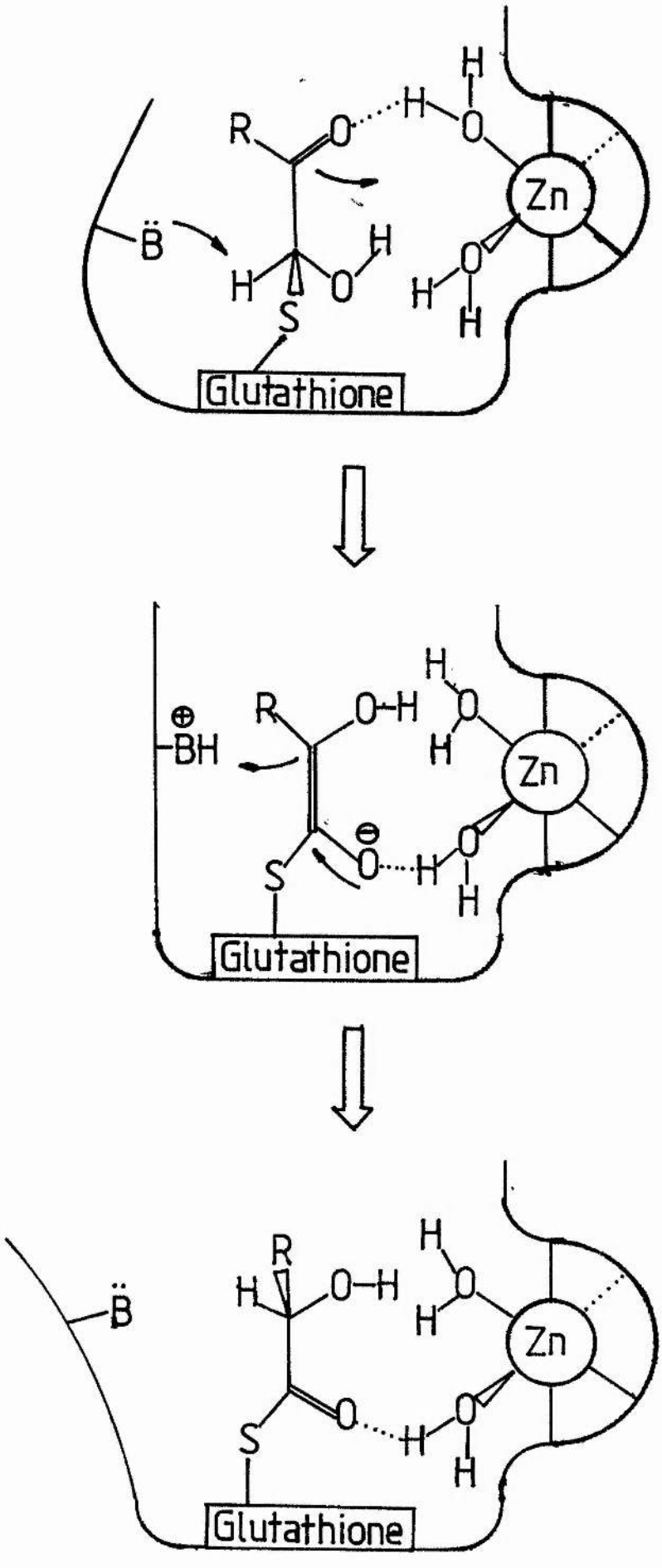

R
O
H
O
Zn
B
H
H
O
S
O
H
H
Glutathione
⊕
BH
R
O-H
H
O
H
Zn
⊖
O
S
O
H
H
Glutathione
R
H
O-H
B
H
O
H
Zn
O
S
O
H
H
Glutathione

Figure 3-1 The Proposed Mechanism Of Glyoxalase I, According To Sellin et al [67].

CHAPTER 4

THEORY AND METHODS

The application of theoretical methods to biologically important molecules has increased dramatically over the last decade. Initially, empirical methods such as Hansch analysis [108] and other Quantitative Structure-Activity Relationships (QSAR) were used, which, although successful in their own right, do not give a clear indication of what is happening at the biological receptor, or enzyme site. Hansch analysis works well if the biological property being investigated relies on gross physical properties, but is not so reliable for very specific bindings, which often occur at physiological sites. Thus, analyses based on a more physicochemical approach were needed.

Theoretical calculations of molecular conformation can provide more subtle information, or can be used as an extension to the familiar Dreiding and CPK models. These models furnish the investigator with conformational information via hard atom contacts, whereas calculations of potential energy surfaces give smoothly varying interactions when neighbouring chemical groups approach one another.

Calculations also provide the means for studying particular molecular properties, especially for isolated molecules. These properties can be relevant to biological reactions. This information is obtained from either semi-empirical or ab initio quantum mechanical calculations, by evaluating the expectation value of the appropriate operator. In principle, any electronic property can be calculated in this manner.

An advantage of these methods over more traditional measurements of gross physical properties, is that the characteristics of parts of molecules, or even molecules which have never been made, can be studied. For example, the electron density on a particular atom or region may be found to be allied to the trend in activity of a particular series of molecules. This could not be deduced by physical experiment, but the electron density is easily calculated by quantum mechanical methods.

The idea of studying molecules which have never been synthesised appeals greatly to drug designers: if a series of molecules seems to exhibit a relationship between biological activity and a particular electronic property, then analogous substituted compounds may be predicted to have certain activities. Unpromising molecules can be eliminated from the list of possible compounds for synthesis, and more likely candidates can be concentrated on. In fact, some synthetic medicinal chemists expect to make only active compounds in the future!

A problem in calculating the electronic properties of molecules in biological systems has been the presentation and subsequent interpretation of the massive amounts of data calculated. This can be alleviated by the modern techniques of computer graphics, as detailed in the next chapter. The properties of simple systems can be studied by considering their symmetry, using group theory techniques, but this soon becomes impracticable. A large, three-dimensional structure requires a four-dimensional method to present results and calculated properties. This can be achieved by the use of colour as one dimension of information and a two-dimensional projection of the other three dimensions (generally molecular coordinates) on the screen plane. To this end, various programs have been devised [109-111], which plot

colour-coded information onto a molecular surface. However, these use expensive, sophisticated equipment, which is not readily available in a typical laboratory. Display devices with fewer bits per pixel available, such as our own Tektronix 4109, are more common, so a method of display on these devices was devised.

The wave function for a particular system is calculated by solving Schrödinger's time-independent equation [112]:

$$\mathbb{H}\Psi_i = E_i\Psi_i$$

where $\mathbb{H}$ is the Hamiltonian operator, a sum of kinetic and potential energy terms. The solutions obtained from this equation, Ψ_i, are eigenfunctions of $\mathbb{H}$ with corresponding eigenvalues, E_i, which are the energies of the system in state i. The equation cannot be solved analytically for any chemical species other than the hydrogen atom, but an approximate solution can be obtained by numerical means, the accuracy achieved depending only on expense in terms of time, money and computer storage space. In the real world of finite resources it is not practicable to calculate exact solutions to the equation, so various approximations are used to reduce the calculation to more manageable proportions.

The main approximations made concern the motions of the electrons in the molecular system. The Born-Oppenheimer approximation considers the motions of the electrons separately from those of the nuclei and the Hartree-Fock self-consistent field (SCF) method considers each electron to be moving in an average field created by the other electrons - the correlated motion of electrons of opposite spin being neglected. Once these approximations have been made, the best single configuration description of the system is the Hartree-Fock limit wave function. To improve on this, one must consider the explicit

correlation of the electrons, but this involves much more complicated and time consuming calculations, which are not possible in the type of work discussed here.

4.1 The Born-Oppenheimer Approximation

The main approximation that is made is the Born-Oppenheimer approximation. This uses the fact that the mass of the nucleus of an atom is much larger than that of an electron, so it moves much slower. The nuclear framework of a molecule can be considered to be fixed, ignoring the kinetic energy terms of the nuclei, and the electrons therefore move in a field created by the fixed nuclear framework. Thus the electronic and nuclear motions can be separated, and the wave function for a particular configuration of the nuclei can be calculated in terms of electronic energy and a constant nuclear repulsion energy.

$$\mathbb{H}^{elec.}\,\Psi_i = E_i^{elec.}\,\Psi_i$$

4.2 The Basis Set

The approximate molecular orbitals are usually constructed from a finite linear combination of atomic orbitals (LCAO):

$$\phi_j = \sum_i c_{ij}\,\chi_i$$

where ϕ_j is the jth molecular orbital and χ_i are the atomic orbitals. C_{ij} are the expansion coefficients defining the molecular orbital. The set of atomic orbitals used to define the molecular

orbitals is called the basis set.

The choice of a basis set is the first step in an ab initio calculation. Although the Slater-type orbitals, which are obtained from atomic wave functions, give good results, most programs employ gaussian orbitals instead. Gaussian functions have a different radial dependence from the Slater-type functions, particularly at the origin and at large distances, but groups of these functions can be fitted to the exponential form of the Slater-type functions. The reason why these gaussian functions are used is that the integrals are much easier to compute using gaussian functions, as the product of two gaussians is a third gaussian, centred between them; this property is used by the integral package of an ab initio program to simplify the complicated many-centre two-electron integrals evaluated in the calculation, as integrals involving four centres and four corresponding gaussian functions can be reduced to two-centre integrals and then to an even simpler expression involving the error function [113].

The fact that about three gaussian functions are needed for every Slater function brings fresh problems in the iterative stages of the self-consistent field procedure, as convergence may be very slow. This is overcome by fixing some of the coefficients relative to others and varying them en bloc, instead of individually. This technique is called contraction of the basis set. The loss of accuracy inherent in this constriction is not very great if the contraction is chosen carefully, whereas the saving in computer time can be.

The simplest type of basis set available is the minimal basis, in which there is one Slater-type function - or set of fitted gaussians - for each orbital. An example of this basis is the STO-3G basis set [114,115] used in some of this work. (Some people may define this as a

sub-minimal basis, as it gives worse energies than a Slater minimal basis set).

The main problem with the minimal basis set is that it is not possible to contract or expand the orbitals to suit their environment, as the exponents are fixed at the start of the calculation. A remedy to this is to use the more complicated split-valence basis sets or the double-zeta basis sets. These are split into two or more parts, allowing the self-consistent field procedure to adjust the size of the orbitals by varying the coefficients of the inner and outer functions. Smaller basis sets of this type, the split-valence sets, such as the 3-21G basis [116,117], only split the valence functions, whereas the larger double zeta basis sets have split core functions too. The 3-21G basis has a contraction of three gaussians for the core orbitals and a split of two inner and one outer contracted gaussians for the valence orbitals. As one might expect, most chemical properties obtained from the split-valence bases are comparable to the double zeta ones, because the behaviour of the valence electrons is much more important than the tightly bound, chemically less interesting core electrons.

A further improvement can be made to the calculation by the addition of more functions to describe each orbital, going to triple or quadruple zeta basis sets. However, it is more effective to incorporate functions of higher angular quantum number in the basis. These are termed polarisation functions, as they allow the electronic charge distribution associated with an atomic orbital to be polarised in response to the electric field generated by the nuclei and electrons. This extra versatility available to the SCF procedure gives good quality results for reasonably sized molecules. An example of this type of basis, used in this work, is the 6-31G** basis [118], which has polarisation functions on all the atoms, including

p-functions on the hydrogen atoms.

The basis set chosen for a calculation is ideally the largest well-balanced one available for the system being studied. The limits of time - both elapsed time and computer time - constrain the choice of basis set, as does the amount of storage space available on the computer. The large inhibitor molecules studied in this thesis, with up to 23 heavy atoms and 10 hydrogens, can only be studied with a minimal basis set on the facilities available in St Andrews. The STO-3G basis set was chosen for this as the limitations of this basis are well documented [119]. The energy calculations on model reactive intermediates were carried out using a split-valence basis set with polarisation d-functions on the sulphur and magnesium atoms, the 3-21G* set [120]. The calculations were also made using the 3-21G basis set [116,117], to study the effect of the d-functions on the system. Single point calculations were carried out at the 3-21G* optimised geometries, using the 6-31G** basis [118], which has polarisation functions on every atom.

4.3 The Hartree-Fock Method

The molecular orbitals constructed from the linear combination of atomic orbitals are in turn used to construct the wave function of the species. The wave function is formed as a determinant - the Slater determinant - so that it is anti-symmetric upon the exchange of any two spatial or spin coordinates; this is a requirement of the postulates of quantum mechanics.

$$\Psi = N!^{-\frac{1}{2}} \begin{vmatrix} \phi_1(1) & \phi_2(1) & \dots & \phi_N(1) \\ \phi_1(2) & \phi_2(2) & \dots & \phi_N(2) \\ \vdots & \vdots & & \vdots \\ \phi_N(N) & \phi_2(N) & \dots & \phi_N(N) \end{vmatrix}$$

$N!^{-1/2}$ is a normalisation factor for the wave function.

As the determinant has a regular structure, it is generally written in a more concise form involving just the diagonal terms:

$$\Psi = | \phi_1 \phi_2 \dots \phi_N |$$

All the calculations in this work have been on closed-shell systems, so the restricted Hartree-Fock method has been used. In this there are pairs of spin-orbitals differing only in their spin - in other words, having identical spatial parts. This gives us a determinant as follows:

$$\Psi = | \phi_1^s \; \overline{\phi_1^s} \; \dots \; \phi_{N/2}^s \; \overline{\phi_{N/2}^s} |$$

in which

$$\phi_i^s = \phi_{2i-1} \alpha$$
$$\overline{\phi_i^s} = \phi_{2i} \beta$$

where α, β are spin functions and ϕ_i^s is the ith spatial orbital.

The energy of the system can be expressed as:

$$E = \int \Psi^* H \Psi \, d\upsilon$$

where $\Psi^* = \Psi$, because we are only interested in real wave functions and we assume that the wave functions are normalised.

This can be expanded to give:

$$E = 2 \sum_i H_i + \sum_{ij} (2J_{ij} - K_{ij})$$

where

$$H_i = \int \phi_i^* H(1) \phi_i \, d\upsilon$$

$$J_{ij} = \int \phi_j^* J_i \phi_j \, d\upsilon$$

$$K_{ij} = \int \phi_j^* K_i \phi_j \, d\upsilon$$

H(1), the core operator, is:

$$H(1) = -\frac{1}{2} \nabla_i^2 - \sum_A \frac{Z_A}{r_{Ai}}$$

This describes the kinetic energy of the electrons and the nuclei-electron attraction terms.

J_j, the coulomb operator is defined by:

$$J_j \phi_i(1) = \int \phi_j(2)^2 \frac{1}{r_{12}} dr_2 \, \phi_i(1)$$

This term is due to the coulombic repulsion between electrons i and j.

K_j, the exchange operator, is defined by:

$$K_j \phi_i(1) = \int \phi_j(2) \phi_i(1) \frac{1}{r_{12}} d\tau_2 \phi_i(1)$$

This term does not have a classical interpretation: it is the exchange repulsion of two electrons of the same spin, a consequence of the anti-symmetry properties of the wave function.

Having defined the trial wave function for the system, we can now use the linear variation method, which is based on the variation theorem, to give us the best set of coefficients defining the wave function. The variation theorem states that the lowest state of a particular symmetry of a system has an energy lower than any other wave function of that system. Thus the coefficients can be changed to find the lowest energy form. The equations used to find this function, the Hartree-Fock equations, are:

$$\left[H(1) + 2\sum_j J_j - \sum_j K_j \right] \phi_i(1) = \epsilon_i^{SCF} \phi_i(1)$$

or

$$F \phi_i = \epsilon_i^{SCF} \phi_i$$

where F is the Fock operator, representing the sum of core, coulomb and exchange operators in the square brackets. ϕ_i are the spin orbitals and ϵ_i^{SCF} are corresponding constants.

As the Fock operator, F, is itself a function of the molecular orbitals, the best molecular orbitals can be found by an iterative procedure in which the trial molecular orbitals are guessed at and used to construct F. The Fock operator is then used to find new molecular orbitals which in turn give a new F. This procedure is repeated until some sort of self-consistency is achieved, that is the ϕ_i produced by F are the same as the previous ϕ_i producing F.

4.4 Properties Obtained From The Wave Function

Once one has obtained the wave function, it is possible to calculate the values of many molecular properties. Most of the properties that we are interested in are one-electron properties and are calculated by evaluating integrals involving the molecular orbitals and the operator corresponding to the property desired. This expectation value of the property defined by operator A is:

$$\begin{aligned} \langle A \rangle &= \int \psi^* /\!/ A \, \psi \, d\upsilon \\ &= \sum_i \int \phi_i /\!/ A \, \phi_i \, d\upsilon \\ &= \sum_i \sum_{jk} c_{ki} c_{ji} \int \chi_j /\!/ A \, \chi_k \, d\upsilon \\ &= \sum_{jk} P_{jk} \int \chi_j /\!/ A \, \chi_k d\upsilon \end{aligned}$$

where P_{jk}, the density matrix, is defined by

$$P_{jk} = \sum_i c_{ki} c_{ji}$$

This density matrix defines the charge density $\rho(r)$ given in terms of the basis set functions used in the calculation.

$$\rho(r) = |\psi|^2$$

Thus, all that is required to specify the property is the density matrix and a set of one-electron integrals,

One-electron properties include the dipole, quadrupole and other higher moments, the molecular electrostatic potential (MEP), the electric field gradient at a nucleus and the force exerted by electrons on a nucleus.

4.5 Energy Calculations

The total electronic energy of the system is obtained from the wave function, as shown above. This in turn gives the total energy of the system - within the Born-Oppenheimer approximation - by addition of the constant nuclear repulsion term. The concomitant orbital energies, ϵ_i, for the spin orbitals, are also directly given as solutions of the eigenvalue equation:

$$\hat{H}\phi_i = \epsilon_i \phi_i$$

The N lowest energies are the energies of the occupied orbitals, whilst the higher energies represent the energy of an electron in the virtual orbital, ϕ_r, using the ground state wave function Ψ_0. This leads us to a discussion of Koopmans' theorem [112]. This states that, assuming the spin orbitals for the ground state and the ionised state are identical, the orbital energy of an occupied orbital, ϵ_i, can be equated to the negative of the ionisation potential for the removal of that electron, and that of a virtual orbital, ϵ_r, can be equated to the negative of the electron affinity for the addition of the electron to that orbital. This neglects the relaxation of the spin orbitals of the N-electron wave function to the optimum N±1-electron wave function and does not include correlation effects either. These correlation errors tend to reduce the error in the ionisation potential value, but increase it in the electron affinity results; therefore electron affinity calculations via Koopmans' theorem are not generally attempted for Hartree-Fock wave functions.

4.6 Population Analysis

An electronic charge assigned to the various atoms in a molecule would be useful to predict the susceptibilities of these to electrophilic or nucleophilic attack. However, although this is an easy concept to grasp, it is not a physical observable. Because it is such a simple notion, there have been various schemes proposed to compute such charges. These give rise to the various population analysis schemes. The commonest of these is the Mulliken population analysis [112]. The electronic population on a particular atom, A, is given by $\sum_i n_i c_{Ai}^2$ for a particular molecular orbital and the overlap population between two atoms, A and B, is given by $\sum_i 2n_i c_{Ai} c_{Bi} \int \chi_A \chi_B d\nu$. The gross atomic population is obtained by assuming that the overlap population can be equally divided between the two atoms, so that the total atomic population is the sum of the net atomic population and half of the overlap populations. To obtain the total atomic charge, the nuclear charge, Z_A, is added to the gross atomic population:

$$q_A = Z_A - \sum_i \left(n_i c_{Ai}^2 + \sum_j n_i c_{Ai} c_{ji} \int \chi_A \chi_j d\nu \right)$$

The numbers obtained by using this process are very dependent on the basis set. That is, if more basis functions are placed on one centre than another, the population results will be distorted, since in the extreme limit, with all basis functions on one centre, all the electronic charge is associated with that one atom. For this reason the Mulliken population values quoted in this work have been used for illustrative purposes in the discussion, to support other observations.

4.7 Molecular Electrostatic Potential

The idea of a charge associated with a particular area of space can be explored by using the electronic density function:

$$\rho(r) = |\psi|^2$$

$\rho(r)$ is the average number of electrons in a unit volume element at point r. This can be localised to a certain region by integrating over that region of space. Thus the net charge on atom A is then:

$$q_A = Z_A - \int_A \rho(r)\, dr$$

where the integration is over the region of space associated with that atom. There is still an arbitrary decision over what this region associated with a given atom might be, as there are no rules defining an atom in a particular molecule. This can be bypassed by forgetting about the constituent atoms of the molecule and concentrating instead on the form of the electron density over the whole molecule, either in a given plane or at particular points in space, such as maxima and minima in the density distribution.

Atomic charges are usually used to investigate the interaction in electrophilic and nucleophilic reactions, assuming that electrophiles will approach the more negative atoms and nucleophiles the more positive. The electrostatic interaction in these cases is obviously the important point. Therefore we are really interested not so much in the charge density as the electrostatic potential which is due to the electrons and nuclei of the system. The calculated atomic charges provide a point-charge representation of this, using a coulombic q/r summation, but the electrostatic potential can in fact be calculated directly from the wave function, without the ambiguities inherent in

the atomic charge calculations. This is defined by the equation [121]:

$$V(r) = \sum_A \frac{Z_A}{|R_A - r|} - \int \frac{\rho(r')\,dr'}{|r' - r|}$$

The first term is the coulombic interaction between the nuclei and a point positive charge at r. The second term is the interaction between the point charge and the electron density function $\rho(r)$. This property is usually depicted as a contour diagram in the plane of the molecule, or in any other plane of interest. The interpretation of the molecular electrostatic potential function has been particularly successful in explaining electrophilic reactions, as the point charge is analogous to a proton.

The molecular electrostatic potential (MEP) is a physical observable, unlike atomic charge: it can be measured in electron scattering experiments, where the incident speed of the electron beam is high [122]. At high collisional speeds, the electronic charge distribution associated with the molecule cannot change fast enough to be perturbed by the approaching point charge (the electron), so the Born-Oppenheimer assumption is a good approximation of the situation. In this case, the electrostatic potential obtained from a molecular orbital calculation, based on the unperturbed charge distribution's interaction with a point charge is a fairly accurate indication of the resultant interaction.

The MEP has been used in many studies of the reactions of molecules in a biological context, notably by Weinstein's group [123], who set out to systematically explore the extents and limitations of the method. They found that small basis sets gave qualitatively the same electrostatic potentials as larger bases [124]; even pseudopotential calculations can give a good description of this

property [125]. However, it is true that the MEP can be basis set-dependent when there are small differences in energy between reactive sites [126], but it is reliable as a qualitative instrument for determining the gross features of the interaction when the interaction is predominantly electrostatic in nature. This interaction can be divided into components by methods such as the Morokuma decomposition analysis of the interaction energy of supermolecules [127]. These components include not just electrostatic terms, but polarisation, charge transfer and exchange energies. Although the latter terms can be important in determining the interaction of molecules at close quarters, the electrostatic repulsions and attractions are generally the dominant contributions at larger distances: therefore the electrostatic potential is often a reliable guide for the initial interactions of molecules, such as which site on a molecule will be attacked by an electrophile.

As a qualitative yardstick, the MEP is useful. However, there are certain qualifications which must be placed on this interpretation: the MEP is dependent on the nuclear geometry of the wave function, that is, it is not perturbed by the point charge. A more realistic measure of the electrophilic reaction would be to calculate the wave function, including the electrophile, at many points on the surface. This is obviously prohibitively expensive for almost all cases, as an SCF calculation is required at every point, whereas only one calculation is needed for the MEP results.

4.8 Geometry Optimisation

The main problem in a large molecule calculation is the quality of the resultant wave function, as this is usually calculated using a minimal basis set. Whilst basis sets such as the STO-3G one [114,115] are fairly good at predicting geometries, the properties and energies calculated can be inaccurate. However, the trends displayed by properties amongst a series of molecules tend to present a more realistic picture. It is of doubtful value to report the minimal basis calculation energies and properties of a single molecule, but this need not stop us from comparing these results to a similar chemical species. For instance, although energy differences do not suffer variational constraints, their error is usually much less than the corresponding molecular total energies, so reaction paths can be studied with some confidence, using quite primitive basis sets. This brings us to the subject of reaction hypersurfaces and geometry optimisation.

The total energies of the system being investigated can be used to study the reaction surface of a group of atoms. At a fairly local level, this can provide information on chemical reactions, whereas, on a global level, it gives details on all the isomeric forms of the constituent atoms. This rapidly becomes complicated with an increasing number of atoms, so most work is done on stationary points and reaction paths on the surface.

Most studies of energy surfaces attempted nowadays use geometrical parameters to define the surface coordinates. This has become common since the methods proposed by Pulay [128] have been included in automatic computer programs, such as GAUSSIAN 80 [129] and GAUSSIAN 82

[130]. Pulay suggested a scheme for determining the forces exerted by the electrons on the nuclei by analytical means. The energy derivatives with respect to the nuclear coordinates are calculated from the wave function. The main advantage of this method, as pointed out by Pulay himself, is that a calculation at one point on the surface yields 3M-6 forces, as opposed to one energy result. (M is the number of atoms). To calculate 3M-6 energies at 3M-6 points on the surface would take considerably longer, for much the same informational value.

There are three geometry optimisation techniques used by GAUSSIAN 82 - the program which I have used the most. These are the Fletcher-Powell [131], Berny [132] and Murtaugh-Sargent [133] algorithms. The Fletcher-Powell procedure uses just the energy and performs numerical differentiations. The other two use analytical gradient calculations, based on Pulay's work, to give information on the slope and curvature of the energy hypersurface. The difference between the two is not in the method of calculation of the forces, but in the determination of the next geometry to be tried. The Berny algorithm uses an initial guess at the second derivative matrix (the Hessian) which is based on empirical values, giving a diagonal, approximate force constant matrix. This works well for simple structures, but does not give a very reliable start for more unusual species, such as cyclic molecules. The first point calculation for these can be improved by directly calculating some of the second derivatives. This is achieved by displacing the relevant geometry variable a small amount and calculating a new energy at this point. This, along with the energy and gradient at the original variable value can be used to calculate the diagonal force constant. The off-diagonal elements of the second derivative matrix can also be calculated if the gradient is calculated at the stepped variable point, but this is obviously slower. Thus, a starting Hessian matrix can be constructed

which is a mixture of empirical and calculated values. It may contain all the elements of the matrix for variables which are thought to be particularly crucial. The best initial Hessian would be a full force constant matrix, obtained by considering every variable in turn, or, as is more likely, from a previous calculation at a lower level.

The optimisation of cyclic geometries is very dependent on the choice of Z-matrix variables, because of the reliance of the constructed Hessian on these variables. This may lead to heavily coupled internal coordinates, unless they are chosen carefully. Schlegel has pointed out [134] that, even in the simple case of a six-membered ring, a sequential list of five bond lengths and 4 bond angles - whilst perfectly well describing any particular planar geometry - affect the sixth assumed bond distance in an additive manner; that is, for relatively small changes in the values of these variables, the remaining atom to atom distance can vary wildly. This does not make matters easy for the optimising procedure. A better geometry can be defined by specifying a cross-ring atom to atom distance and then constructing the other four atomic positions with independent angles to this first variable.

If the Berny procedure still fails to find a stationary point on the surface, then the Murtaugh-Sargent method can be tried. This uses no empirical values for the Hessian and starts with a simple unit matrix. The method is much slower than the Berny one, but it will usually be more stable (that is, the energy values should improve each time).

Almost all the inhibitory molecules studied have had some form of cyclic structure. It was very difficult to find an optimised, stationary point for many of them, especially coumarin.

4.9 The Strategy Of The Calculations

The aim of the calculations was to provide information on both the reactive mechanism of the glyoxalase I enzyme and on the inhibitors of it. These two topics require different approaches: the mechanism has been investigated by using a model reaction scheme of the reactive route via the proposed reactive intermediates. This required a good calculation, including polarisation functions on the second row atoms. The inhibitor calculations could not be carried out using anything larger than a minimal basis set, as there are too many atoms in them. The examination of the reaction scheme requires full optimised geometry calculations at every point of interest on the surface. The optimisation of the geometries of the inhibitor molecules could not be realised totally: the basic skeleton of these was optimised, but the substituent groups could not be varied along with the skeleton in every case, because of the time limits imposed on the work. This is not as drastic a stricture as it at first seems, as it was found that the form of the MEP was not much affected by the structure of the compound, once the forces calculated to be on the atoms was less than 0.04 a.u.[135]. The minima in the MEP occur at the same places, but vary in depth, depending on the wave function's quality, but qualitatively the MEP has the same form.

In this work there is very little information available on the nature of the active site structure. The excellent work of Mannervik's group has provided some pointers to the probable coordination around the zinc atom of the active site [67-69], but no X-ray crystal structure of the enzyme has been reported. It was not possible to carry out any supermolecule calculations involving the model reaction intermediates or the inhibitors with active site residues, as these residues and their relative positions are not known. A more qualitative approach, in which the MEP of the inhibitors were compared to each other, or to the model scheme, was tried.

As glyoxalase I is such an efficient enzyme [136], it seems reasonable - if we assume that the interaction is largely electrostatic in nature - that the binding of the substrate would occur at a template region of the active site, in which the electrostatic potential was complementary to that of the substrate [137]. Therefore, it may be possible to predict the nature of the active site by comparing the data obtained from model intermediate and inhibitor calculations. As it is difficult to directly compare a neutral inhibitor molecule with an anionic species for which the electrostatic potential is negative outside the van der Waals surface in all directions, the information gleaned from the inhibitors must be processed separately from that for the model reaction, or compared only to the neutral intermediates. By these comparisons, it was hoped that more powerful inhibitors could be devised.

CHAPTER 5

GRAPHICAL DISPLAY PROGRAM

The study of the inhibitor molecules in this work was based largely on the examination of the molecular electrostatic potential surrounding them. The method adopted in the earliest studies consisted of calculating the molecular electrostatic potential on a grid of points defining a planar section through the molecule. This used a modified GAUSSIAN 80 program [138] which includes the DENPOT program [139] in overlay 6. Thus contour plots of V(r) in the molecular plane could be drawn. It soon became obvious that a more useful depiction of the electrostatic potential due to the molecules could be created if a three-dimensional representation was available. This representation is even more essential for non-planar molecules.

To display a property defined by a three-dimensional coordinate system, instead of a two-dimensional grid, requires an identifiable fourth value or property which can be associated with any given point on the surface of the three-dimensional object. The contour lines of a planar section become sheets in three dimensions and these cannot just be labelled with an appropriate number, as they are on the planar contour maps, as this gives too complicated a diagram to be of any use. The answer to this is to use colour coding instead of labelling on the diagram. The use of colour can be thought of as adding an extra dimension to any n-dimensional representation of information. For instance, the depiction of a property, p, based on two coordinates, x and y, can be displayed either as a colour-coded contour map of exact line values or as a colour-filled plot of all values of p between two border values. Each (x,y) point in the representation can have a colour associated with it which corresponds to the value of p at that point. Both of these techniques have been used on geographical surface contour maps. Likewise, a three-dimensional surface can be

colour-coded to provide more information, either as coloured contour sheets, or as a colour-coded surface. Both of these representations have been used to depict properties of molecules on a colour terminal screen [109-111]

Before the reasons behind the particular method that has been chosen can be presented, it is necessary to list the resources available and the limitations imposed on the program by these resources: the devices at our disposal for the purpose of displaying the properties are a Tektronix 4170 graphics coprocessor and a Tektronix 4109 terminal, coupled to a Tektronix 4695 colour plotter. This forms a self-contained system for graphics work, as the 4170 runs CP/M-86, a version of the widely used CP/M operating system, allowing the complete manipulation - from writing to running - of FORTRAN-86 programs. FORTRAN-86 is a version of FORTRAN designed to run on an 8086 microprocessor. A set of FORTRAN-callable subroutines, the DTI (Direct Terminal Interface) package, can be used for graphics purposes; these correspond to each of the terminal's display commands, giving an easy interface to the terminal's features.

The terminal is a raster device with a 4-bit frame buffer. This means that each point, or pixel, on the screen is accessed sequentially, like a television, allowing areas of colour to be produced. This is not possible with a vector display, which can only draw and move from a starting to a finishing point, like a pen plotter. This ability to display solid colours allows realistic representations of three-dimensional objects to be produced. More sophisticated (and expensive) raster terminals have more bit-planes than the 4109. This allows the display of more colours than the sixteen available to us, as a raster device with n bit-planes can display 2^n colours at one time. Thus, the effects of light, such as

highlighting and shading can be depicted to give even more realistic pictures. This is not feasible on our terminal, unless the picture is monochromatic.

The 4109 also has good picture manipulation in two dimensions via the segmentation properties of the terminal, so that pictures can be scaled and transformed, or stored in terminal memory and re-drawn at a later stage. Interactive communication with the terminal is provided through a joy-disc-controlled cursor, which can pick items from the screen.

We wished to use all these features in a simple interactive program, allowing the display of molecular properties in an easily understandable format, and the subsequent manipulation of the molecules on the screen, including scaling and rotation.

As mentioned before, there are two possible ways of representing molecular properties on the graphics terminal: as contour sheeting or as a coded surface with areas of colour where the various contour sheets intersect the surface. The coloured sheeting method has been used in the ChemGraf program [109] to represent the charge distribution around a molecule. It is very difficult to interpret, especially if more than about three contour levels are used. Unless the viewer has a good visualisation of three-dimensional objects, the resulting plots will not provide much useful information without considerable mental effort on his part. The second method, of a colour-coded surface has also been used in ChemGraf: a colour-coded chicken-wire mesh can be constructed around a molecule's van der Waals surface. This is also difficult to interpret, as the program draws the whole of the chicken-wire mesh, both back and front. This can lead to optical effects similar to the Necker cube illusion [140], in which the viewer

is not certain which is the back and which is the front of the constructed surface. Again, considerable mental effort is required to extract information from the plot. The confusion over which surface is which on the chicken-wire mesh could be remedied by implementing a hidden-line removal algorithm which considered the polygons formed by the mesh lines to be opaque, eliminating the rear lines of the mesh.

The main feature of a raster terminal is its ability to fill areas with a solid colour or pattern. This is because each addressable point, or pixel, in a rectangular matrix, or raster, covering the screen is assigned a colour. This colour can form part of a picture, or can be the background colour. In contrast, a vector device does not store details of the whole screen, point by point, it stores a linear list of locations which define the picture in terms of move and draw commands, like a pen plotter. The ability of a raster device to fill areas with colour has been utilised in the graphics program. As the representations obtained using meshes were not satisfactory, the use of colour could improve the display by making the polygons defined by the mesh framework opaque, obscuring the back of the object. If the picture is constructed not out of lines, but out of the polygon surfaces between the lines, then a solid, structured image is the result. Also, if the polygons are drawn in order from the rear, then the polygons nearest the observer will obscure the areas at the back, avoiding the illusions created by the mesh representation. This just involves the computing time necessary for a depth priority sort, that is an ordering of the polygons depending on the z-coordinate; there is no need to indulge in complicated and time-consuming hidden-surface calculations. Thus, it was decided to write a program which drew surfaces constructed of colour-coded polygons. This could depict a property on the van der Waals surface of a molecule by a coloured dot on the surface at the point that the property had been calculated. If

these dots were sufficiently close together, then they would form a polygonal mesh, obscuring the back of the molecule as desired. Alternatively, the back of the molecule could be hidden by using a coarser mesh of points, but making the representative dots larger, until they adjoin each other. This leads to a slight distortion of the information presented, as the colour of the dot only represents the calculated property at its centre, but a compromise can be reached between the size of the dot and the calculation of properties without too much loss of detail. To give a better idea of the shape of the surface, the dots could be replaced by discs tangential to the surface at that point, but, as a raster terminal has a finite number of addressable points, circular discs cannot be drawn properly, so it was decided to use filled polygons instead of discs. Two types of polygon were chosen: a simple square, for rapid construction of the surface, and an octagon, which is fairly similar in shape to a disc, for final pictures.

The method chosen was to depict the property at any point (x,y,z) by a coloured polygon, normal to the vector (xn yn zn). The size of the picture could be changed, as could the size of the polygons, allowing the amount of area these cover on the molecular surface to be altered. This method is facilitated by the fact that a copy of the MS program written by Connolly [141] is available on the St Andrews VAX. This outputs coordinates and unit vectors perpendicular to the surface calculated. This Connolly surface is calculated by rolling a probe sphere around the atoms concerned to give a solvent-accessible volume, or a van der Waals surface if the probe radius is zero [142]. The coordinates can be input to a modified GAUSSIAN 80 [143], to calculate the electrostatic potential at those points and the results combined to produce the datafile for the display program.

5.1 Program Method

The basic idea of the program transforms is to align a square (or other polygon), centred at the origin and parallel to the z=0 plane so that it is at (x,y,z) and perpendicular to a vector (a b c). The square is then transformed by a perspective transform, to give a more realistic picture and colour-coded, according to the contouring.

The transformation procedure is given below:

Consider a square centred at (0,0,0) and in the z=0 plane. We must rotate about the x and y axes to map the unit vector z=1 onto the unit vector (a b c) and then translate it to (x,y,z).

The rotation matrices are:

$$R_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & c/v & -b/v & 0 \\ 0 & b/v & c/v & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$R_2 = \begin{bmatrix} v & 0 & -a & 0 \\ 0 & 1 & 0 & 0 \\ a & 0 & v & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$v = (b^2 + c^2)^{1/2}$$

R_1 is a rotation of the unit vector z=1 by ϕ about the y axis to give the x-component, a. The resultant unit vector is rotated by θ about the x-axis to give the y- and z-components, b and c. This rotates the z=1 vector so that it is parallel to (a,b,c).

To translate the vector to (x,y,z), the matrix is:

$$T_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ x & y & z & 1 \end{bmatrix}$$

Therefore the full transformation is:

$$[X' \; Y' \; Z' \; 1] = [X \; Y \; Z \; 1] \; R_1 \; R_2 \; T_1$$

$$= [X \; Y \; Z \; 1] \begin{bmatrix} v & -ab/v & -ac/v & 0 \\ 0 & c/v & -b/v & 0 \\ a & b & c & 0 \\ x & y & z & 1 \end{bmatrix}$$

So,

$$X' = v X + a Z + x$$

$$Y' = -ab/v \; X + c/v \; Y + b Z + y$$

$$Z' = -ac/v \; X - b/v \; Y + c Z + z$$

Now, for a square 100 units by 100 units, positioned at (-50,-50); (50,-50); (50,50); (-50,50), in the Z=0 plane, the equations become:

$$X' = v\,X + x$$

$$Y' = -ab/v\,X + c/v\,Y + y$$

$$Z' = -ac/v\,X - b/v\,Y + z$$

where

X = ± 50, the width of the original square being 100,

Y = ± 50, the height of the original square being 100,

(x,y,z) are the coordinates of the centre of the transformed square,

$\underline{(a\ b\ c)}$ is the vector normal to the surface at that position,

$v = (b^2 + c^2)^{1/2}$

These equations are used on the points above using the normal vectors and coordinates from the MS program. The X' and Y' points are plotted on the display, ignoring the corresponding Z' points if an orthogonal projection onto the screen plane is required, but the Z' information can be utilised to give a perspective projection as elucidated in the next section.

5.1.1 Perspective Transform

The representation at the moment is an orthogonal projection, that is, it takes no account of the distance from the viewpoint to the object: parallel lines do not converge to vanishing points. However perspective depth can be achieved by another simple transformation. The perspective picture is obtained by projecting the object points onto the display screen plane [144].

The transformation is:

$$X' = \frac{D\,X}{S\,Z} \qquad Y' = \frac{D\,Y}{S\,Z}$$

where D is the distance from eye to screen

S is the height of the screen

This uses the points obtained from the rotation and translation transformations to give a perspective representation. This changes the polygons, making them larger when they are nearer to the observer, amplifying the curvature of the structure, but can also lead to distorted pictures if too small a D/S ratio is used.

5.1.2 Hidden Surface Removal

The problem of removing hidden surfaces from the image is an important one in many areas of computer graphics. There have been many methods evolved for dealing with this. (See, for example "Procedural Elements For Computer Graphics" [145]). However we can circumvent this by using convenient features of the raster terminal: because the terminal can fill areas with colour, polygons in front will obscure those drawn behind if a simple ordering of the points by z-coordinate size is made, using a suitable sorting routine. If we draw the furthest polygons first, and then draw progressively nearer ones, the front ones will be drawn over and obscure the others. Provided the polygons are large enough to obscure most of the ones which would not be seen in a totally solid representation (when the polygons are so close together that the surface is effectively continuous), this method works well enough. There does need to be some gap between polygons, though, because large overlapping polygons do not give the impression of curvature to the surface as the borders and hence the orientation of the polygons cannot be seen. A further "trick" incorporated into the program is to draw all the polygons which face away from the observer. These, evidently, cannot be seen, but if they are coloured identically to the background, they can "rub out" polygons behind them, giving a better picture. They will not obscure every one of these hidden polygons, but are likely to remove twice as many as previously was possible; the drawback is that it takes twice as long (on average, with convex surfaces) as when just the forward ones are drawn.

5.2 Rendering

Rendering is the process of producing realistic images on a display screen. This includes depicting the effects of illumination on the object, the texture of the object surface, the transparency of the object medium and the shadows thrown by the object on a background. These can all be introduced into an image on a sophisticated display, but on our simpler device the effects may be beyond the capabilities of the terminal, or may be irrelevant. The features which can be ignored totally are texture and transparency, as our image polygons can be considered to be perfectly opaque with a smooth planar reflecting surface. The shadows produced on parts of the molecule by other parts could be considered, but the idea of discrete points on the surface would be compromised, as some comparison of polygons along ray lines would be needed. In fact, this is exactly analogous to the hidden surface problem, but along the light vector instead of the screen plane normal. These would both require the comparison of all polygons with each other and a complex clipping algorithm would be needed to take care of partially obscured or shaded objects. The time spent in this comparison would not justify the improvement in the image for the terminal being used, especially as the molecules studied do not have convoluted concave regions, as the atoms are considered to be spherical.

A detail which has been used to enhance the display is the addition of shading. As mentioned above, the limitations of a four bit-plane display prevent continuous shading models and other special effects, but the angle of a particular polygon relative to a certain light vector can be used to calculate whether rays from a direct light source will reach it. If not, the polygon can be coloured grey. This

is achieved by considering the dot product of the polygon normal vector and a light vector. If the angle between them is less than 90° or more than 270°, there should be no shadow. Therefore, if cos θ is no greater than 0, there will be no shadow. The dot product of the two vectors provides this information, since the magnitudes of the vectors are always positive so the only negative dot product will occur when the angle is less than 90 or greater than 270 degrees. If no shading of this kind is desired (the contouring information is destroyed by colouring the surface grey) then the light vector can be made parallel to the viewing vector, so no shading is visible.

5.3 Program Operation

The data for the program input is obtained in the following manner: once the geometry of the molecule of interest has been decided on, the MS program written by Connolly [141] is used to calculate the coordinates of the surface of interest. These points are used as input for the modified GAUSSIAN 80 program [143], which outputs a file containing coordinates and the associated electrostatic potential value at that point. This file is combined with the MS program output to give a data file containing (x,y,z) coordinates, the associated (xn yn zn) vector and a number to represent the colour index at that point, calculated by using a list of contour values. This pre-processed file is then transferred to the Tektronix 4170 from the St Andrews VAX computer, in order to run the graphics program.

The input parameters for the graphics program are prompted for by a question and answer, free format session once the program is executing. If the user is a novice, he can ask for additional prompting with examples of typical input values. The graphical data is stored in a file containing each point, its corresponding normal vector and a colour index for the contouring.

After the name of the file has been established, the program asks for details of how the picture must be presented. The user is asked for rotation parameters, in terms of rotation about x- y- and z-axes, and scaling parameters for the polygons and also for the whole picture. This last gives a means of varying the size of the polygons relative to the molecule itself. This allows the user to cover more or less of the surface as desired. The scaling of the picture is given as a screen quotient, S, which is the ratio of the height of the viewing area to the picture, so the larger the picture desired, the smaller the S value. This parameter is linked to the perspective transform via the D/S ratio, D being the distance from eye to screen, which is also prompted for. Both of these values are in arbitrary units, as it is the ratio D/S - a dimensionless quantity - which is used in the transform.

The above parameters set up the gross picture in the viewport allocated for the picture on the screen, but further information is needed on the style and storage of this image: the user is asked whether he wants to use a square or octagon representation for his data points - the square version is quicker, but does not give as good a picture as the octagon version. He is also asked whether he wants to draw the polygons facing away from him, to try to obscure parts of the picture which should be hidden. This is not always necessary. The

light vector for the shading is also required from the operator; this is given as three real numbers in the x- y- and z-axis directions. Finally, details for terminal storage are requested. The segment number and view number are prompted for. The terminal stores pictures in views, which are collections of picture segments at locations on the screen known as viewports. Images can be stored in these views and later recalled for comparison.

Once these parameters have been entered, the program calculates and displays the picture in a large viewport covering the top left of the screen. Any of the parameters mentioned above can be changed by using an interactive, joy-disc-driven cursor, which picks the relevant parameter from a menu displayed to the right of the picture. A new value is prompted for and the menu is renewed accordingly. Once the user is satisfied that he has altered all the parameters necessary, the picture can be re-drawn, by exiting the menu. This is a perfectly general manipulation of all parameters, from the data file to the storage segment being used, so different views of the same molecule, or of different molecules can be produced with one execution of the program.

5.4 The VAX Version

The above program runs on the Tektronix 4170 graphics co-processor, using the DTI graphics routines at a baud rate of 19,200 baud. A version of the program has also been written for the St Andrews VAX 11-785 computers. As there were no graphics sub-routine libraries available on the VAX, the graphics primitives had to be written to create an equivalent DTI library. This was achieved by

writing sub-routines which use the Tektronix 4109 terminal's escape sequences. This involves converting parameters into a host syntax form. Different conversion procedures are necessary for the conversion of characters, character strings, small integers (0-127), integers, reals, integer arrays, real arrays and xy-coordinates on the screen [146]. The converted host syntax escape sequence must then be sent to the terminal. The easiest way that I have found to do this is to use an ordinary FORTRAN WRITE statement, which is interpreted by the terminal and ignored by the host, resulting in the desired graphics command. For instance: <ESC>MP% where <ESC> represents the character for the escape function, i.e CHAR(27), is the command to change the colour index associated with filling a polygon to 5, as MP is the code for SELECT-FILL-PATTERN and % is the code for -5.

There is one more problem which must be dealt with: some escape sequences, known as terminal report request items, result in a coded response from the terminal to the host, for instance a REPORT SEGMENT STATUS escape sequence will result in the terminal sending an encoded report of segment characteristics to the host. (A segment is a collection of graphics primitives which can be manipulated en bloc). If these reports are not trapped and dealt with properly, then the host will hang the program, or crash it, as it cannot decipher the report. The facility which is used to remedy this is the $QIOW Queue I/O Request And Wait For Event Flag system service [147]. This has been used to enable the interactive features of the program - the picking of items with the cursor - to be kept in the VAX version. The routine used is a modified version of one written by Mr J. Crowe of the St Andrews University Computing Laboratory [148].

5.5 Improvements

The program was written for a specific purpose: the display of properties on the van der Waals (or Connolly surface) of a group of similar molecules. It was not necessary to make it too sophisticated, but the framework is there for any future development work.

There are some limitations of the program which have become apparent with use: the input is restricted and the program is slow at lower baud rates, but the presentation of data is good, considering the limitations of the terminal display.

The main criticism that can be levelled at the input is that it is in a pre-processed form, consisting of scaled coordinates and a colour index number for each point. This was because the VAX is much quicker at number-crunching than the 4170, so as much of the processing as was possible was done on the VAX, prior to transfer to the graphics co-processor. This means that certain parts of the display are not flexible: to change the contouring, one must run the conversion programs on the VAX again, to process the original data file, giving a new input file for the Tektronix. It might be worth making the program input more general, so that the original data files from the GAUSSIAN 80 program [143] could be processed directly.

If the actual coordinate and associated property data were used by the program, it would be possible to consider the scaling of the picture in a more quantitative fashion, instead of as a ratio to the viewport size. This was not important in this thesis, because the two main groups of molecules that were of interest, that is, the flavones and the coumarins, were of similar size within their respective groups,

so one picture could easily be compared with another. However, if a range of molecules of varying size were to be investigated, the adjusting of the size of the picture relative to the viewport instead of to an absolute value could be distracting, as smaller molecules appear to be as big as the larger ones. On the other hand, if the molecules being studied are so dissimilar, it is perhaps not such a good idea to compare and contrast them by this method in the first place.

In order to have an absolute scaling scheme, in which the picture is drawn subject to a scale of say, three Angstroms per thousand screen locations, then the picture would have to be drawn using an orthogonal projection, with no consideration of the perspective, as this distorts the coordinate system. Thus, the incorporation of a measurable scaling results in some loss of shape in the image, as the distortion produced by a perspective transform aids the eye in interpreting the image's shape.

Another improvement which could be made in the program is to have the option of changing the orientation of the image relative to the last view drawn, instead of relative to the original data. This method is used in ChemGraf [109] to rotate any molecule drawn on the screen and is useful for small adjustments in the orientation of the image. It does have its disadvantages, though: in order to re-create a picture it is necessary to go through all the sequential rotation transforms in the right order, or to keep a running tally of them, whereas if an absolute transformation of the original coordinates is used, it is only necessary to keep a note of the one transform which resulted in the desired picture. It would be a trivial matter to include a relative/absolute mode toggle in the program options, as it would just require a flag to control whether the transformation subroutine,

TRANSF, operated on the original coordinate array, COORD, or on the previously transformed array, TRANS.

The speed of program operation is satisfactory on the Tektronix 4170 graphics co-processor, which runs at a baud rate of 19,200 baud, but it is a serious problem on the St Andrews VAX computers, which run at 1200 or 2400 baud; it can take twenty minutes to draw a complete picture. This makes a mockery of the intention to create an interactive, user-friendly program, as the timescale is closer to that of a batch job. Any increase of speed that can be achieved is obviously desirable. One way of achieving this would be to concatenate groups of graphics primitives into one command string, as this cuts down the number of host to terminal messages sent. The best candidate for this approach would be the group of SELECT-FILL-PATTERN, BEGIN-PANEL, MOVE, DRAW and END-PANEL commands necessary to construct each polygon. As the drawing of these is the major time factor in the program operation, any saving of time in their construction would make a noticeable improvement.

CHAPTER 6

MODEL SCHEMES FOR THE ENZYMIC ACTION

The glyoxalase I catalysed reaction consists of two proton transfers or two proton abstractions and two protonations. There are various pathways that can be devised, based on the exact order in which these abstractions and protonations can take place. Lavery and Pullman [149] proposed a series of single proton transfer reactions between a model substrate and a model product molecule. The methylglyoxal reaction was used and glutathione was replaced by methanethiol. The calculations were minimal basis set ones, using a standard set of bond lengths and angles, with limited minimisation of C-C, C-O and C-S bonds. The oxygen atoms were maintained in a cis configuration, as the glyoxalase reaction seems to involve this conformation (all the inhibitor species have cis oxygens). Calculations involving a chelated magnesium ion were also carried out. The scheme adopted by Lavery and Pullman is shown in Figure 6-1.

The authors found that the pathway A->C->E->F->H was the most likely one, for both the reaction with and the reaction without a magnesium ion. However, they could not rule out an alternative path via a doubly deprotonated dianion, D. The first-mentioned pathway consists of successive deprotonations and protonations. It also includes the much-discussed enediol "transition state", which is actually a reactive intermediate species, as are all the molecules B to G.

This study seemed a good place to start the investigation of the glyoxalase I mechanism. Two obvious improvements were immediately evident: a complete geometry optimisation at each step would give more trustworthy results, and a more extensive basis set would allow more subtle effects, such as the role of the sulphur atom in the reaction,

to be examined.

The basis set chosen was the 3-21G* one [118], as explained in the next section. This basis is the same as the well-known 3-21G set [114,115] for hydrogen, carbon and oxygen, but includes d-functions on sulphur and magnesium.

The results are listed in Tables 6-2 to 6-10 and the optimised geometries are shown in Figures 6-2 to 6-10. The most notable factor, in a comparison with the work of Lavery and Pullman, is the absence of data for the molecules B, D and G. On full optimisation, at the RHF/3-21G* level, these three molecules proved to be unstable, flying apart to form groups of smaller molecules.

The instability of molecule D is perhaps not surprising, considering the double negative charge on such a small species, but it is nevertheless a very important result: this is because the only available pathway is now via C, E and F, as found by Lavery and Pullman. We do not have the dilemma faced by these authors in choosing between two energetically similar routes, as it is impossible for the reaction to proceed via molecule D.

6.1 Choice Of Basis Set

A large basis set with many polarisation functions is required in order to approach the Hartree-Fock limit for even a small molecular calculation. Unfortunately, the constraints of time and computer resources, such as CPU time and disk space allocations, severely limit the size of basis set available. This has led to the development of

basis sets for the study of different aspects of computational problems. For instance, the MINI and MIDI basis sets of Huzinaga [150] were designed to give good energy results, comparable to much larger basis sets, whereas the split-valence basis sets, such as 3-21G [114,115] give geometry results comparable to larger sets. Therefore, it should be possible to find a basis set for most computational problems in chemistry.

One of the main aspects of this study is the presence of a sulphur atom in the compounds, derived from the thionyl group of the original glutathione co-enzyme. To take this into account properly it was decided to use a basis set with 3d-functions, as sulphur has low-lying, empty d-orbitals which can take part in the bonding and electronic properties of its compounds.

The sulphur atom of the glutathionyl moiety could stabilise the negative charge on the intermediate anionic species to help the enzymic reaction. Certainly, some of the model reaction schemes that have been studied seem to require the presence of glutathione, in order to produce an analogous reaction to that of glyoxalase I. These include the work of Hall et al. [78], in which they proposed that the reaction of glutathione with the aldehyde, prior to the catalytic action of glyoxalase I, converted the aldehydic proton into an acidic one. This acidic proton would be stabilised by the adjacent sulphur atom.

There has been much controversy over the role of d-orbitals in carbon-sulphur compounds over the last two decades. In a series of papers [151-155], Oae and others invoked 3d-orbital resonance to explain the acidity of a hydrogen atom adjacent to a sulphide group. They stated that, although the 3d-orbitals of the sulphur atom in the atomic state were too diffuse to overlap effectively with the

2p-orbitals of a carbon atom in neutral molecules, 3d participation did seem to occur when the sulphur was attached to a carbanion or an unsaturated group.

Later work seemed to show that the presence of the d-orbitals on the sulphur atom did not affect the molecule in any way; the lowering of the energy on addition of the 3d-orbitals was just due to an increase in the flexibility in the basis set. The acidity of protons attached to carbons next to a sulphur atom was attributed to the stabilisation of the anion by polarisation of the sulphur [155] and the coefficient matrix did not show any participation of d-orbitals in the higher occupied molecular orbitals [156].

The calculations by Wolfe, Rauk and Csizmadia mentioned above [156] used a d-function exponent of 5.42 in the basis set. An objection could be raised that this exponent would result in very little overlap, restricting efficient bonding. This point was discussed in earlier work by two of the authors [157] and was investigated further for the case of hydrogen sulphide. The optimised exponent of 5.42 (for atomic sulphur) was replaced by one of 0.25, which, according to the authors, gives a maximum in the d-gaussian function at about two bohr from the nucleus. The resulting character of the molecular orbital contribution from the d-functions was still of s-type, being a linear combination of d_{xx}, d_{yy} and d_{zz}, with negligible contribution from the other d-functions, implying no appreciable participation of d-orbitals in the bonding. The authors stated that they expected that a near-Hartree-Fock basis would reduce the contribution of the xz and yz functions even further, as the d-gaussian component would be over-emphasised in their small basis set.

The results obtained by Rauk and Csizmadia could be attributed to a bad choice of exponent for the d-functions, nevertheless. Their first choice of 5.42 was obtained by optimising the exponent in hydrogen sulphide by energy. An experimental geometry was used for the molecule. The exponent does seem to be exceedingly large - the 6-31G* basis set and the 3-21G* basis derived from it have an exponent of 0.65. The resulting maximum value of d_{zz} is at only 0.43 bohr using this exponent. The authors investigated the wave function using an exponent of 0.25, which corresponds to a maximum at 2.0 bohr; this is probably taking the problem to the other extreme, giving too diffuse a function for good bonding in a neutral molecule. Therefore it would be fair to state that the resulting coefficients of the wave function are not that reliable.

There has recently been a re-examination of the earlier work on sulphur-containing carbanions. The use of improved methods has revealed that sulphur 3d-orbitals do play a part in the bonding molecular orbitals of these species [158]. Full geometry optimisation of the molecules being studied revealed that the previously used C-S bond length of 1.819 A (the experimental length) did not allow the d-orbitals to exert their influence fully as the distance was too great. Geometry optimisation of the CH_2SH anion at the 3-21G* level proved that the C-S bond length shortened relative to the 3-21G optimised structure. This shortening was now recognised to be due to conjugative stabilisation of the carbanion by the SH group.

Thus, the presently accepted view, on calculations involving sulphur-carbon bonds, especially when carbanions are involved, is that full geometry optimisations are imperative. A proper description of molecular properties require the inclusion of d-functions in the basis

set, too.

The system that is at our disposal generally has less than 100,000 blocks of disk space available on it and this storage space can fluctuate widely. Therefore, for the study of compounds with at least two carbons, two oxygens and a sulphur atom, it was deemed necessary to use the split-valence 3-21G* basis set [118], which includes d-functions on the sulphur atom. This basis set and the closely related 3-21G set [114,115] have been shown to give comparable geometry results to larger basis sets, such as the 6-31G series [159].

The 3-21G* basis set has been the standard basis for the model scheme calculations in this work, but latterly a grant of resources on the CYBER 205 at the University of Manchester Regional Computing Centre has allowed the extension of the investigation to the larger 6-31G** basis. The restrictions imposed upon these new resources have prevented new optimisations of the species, but 6-31G** calculations at optimised 3-21G* geometries have been made.

The basic model scheme that was described above has been modified by various substituents to gain an insight into the mechanism of the reaction. The substitutions effected have been replacement of the thiol group by a hydroxy group or a proton and extension of the thiol into a methyl sulphide group, to more properly mimic the glutathione coenzyme. (This latter change was also made possible by an increase in computer resources).

The hydroxy- and proton-replaced molecules were optimised using the 3-21G basis set, as this is identical to the 3-21G* basis for these compounds, which have no second-row elements. The results are listed in Tables 6-2 to 6-10.

6.2 The Overall Reaction Scheme

The total energies of the species discussed below are given in Table 6-1. An initial discussion of the overall reaction scheme is given here.

The thiol-substituted model scheme - that is the original scheme - does not feature such large energy differences between the intermediates as the hydroxy- and hydrogen-substituted schemes. The overall reaction shows a stabilisation of -0.0138 a.u., compared with -0.0327 a.u. for the hydroxy-substituted scheme and 0.0 a.u. for the hydrogen-substituted scheme, in which A is the same molecule as H and C is identical to F. Thus, in terms of the overall reaction, it would profit the enzyme to use an oxygen-based substrate instead of a sulphur-based one, as this is the most energetically favourable reaction. However, if one looks at the energy differences along the proposed pathway, it is obvious that there is a substantial gain to be achieved in using sulphur, in that the activation energies needed to remove the protons from the molecule are less severe in the case of the thiol-substituted A->C and E->F reactions than in either of the two analogous schemes.

The choice of isolated molecules, neglecting solvent effects, is justified because this particular enzyme must have a very well protected active site. This is shown by the difficulty in observing solvent protonation when the reaction was carried out in D_2O [73,74] or tritium-enriched water [75], which led to the belief that the enzyme catalysed a reaction analogous to the Cannizzaro reaction. The proton transfer is so well shielded that the solvent does not play a great part in the reaction, as shown by the work of Hall et al. [78,79],

where only a small amount of solvent incorporation was detected by nmr.

Thus we do not suffer from the great bugbear of biological calculations: solvation effects. This topic has an undeserved reputation as the sticking point of many theoretical simulations of chemical reactions; this does not apply exactly to the consideration of enzyme sites, as these often have highly polarised regions not similar to a wet chemistry environment, and the only way to treat this situation, in the absence of detailed information of the active site structure, is a careful examination of the properties of the various substrates and inhibitors. This is attempted in the following chapters.

6.3 The Substrate Model: Molecule A

This particular species is a model for the glutathione-aldehyde adduct, the hemithioacetal used by glyoxalase I. The chemical groups not joined to the oxygen atoms have been replaced by hydrogen because of computer disk space limitations. This enabled the utilisation of a split-valence basis set with d-orbitals on the sulphur atom as discussed above. The groups replaced by a hydrogen atom were the alkyl or phenyl group of the aldehyde and the glutathione tripeptide, except for a residual thiol group.

Molecule A has been studied at both the 3-21G and 3-21G* level, in order to examine the effect of the d-orbitals of the sulphur upon the bonding and chemical properties. Full geometry optimisations, using both these basis sets have been made. The resulting geometries are

listed in Table 6-2 and the 3-21G* geometry is shown in Figure 6-2.

The largest difference between the two geometries is a 4% shortening of the C-S bond length upon the addition of d-orbitals to the sulphur basis set. This immediately suggests that there is some (d-p) π bonding between these two atoms. Other large effects are the widening of the O3-C1-C2-H7 dihedral angle (2.7%) and the O3-C1-C2-S5 angle (1.0%), such that the sulphur approaches the O3-C1-C2 plane, pushing the hydrogen further away from it. The sulphur hydrogen bond length is also decreased by 1.8%. The C1-C2-S5 bond angle is increased by 1.6% and the C1-C2-H7 bond angle is decreased by the same amount. The rest of the molecule is affected by less than 1% changes. Thus, as would be anticipated, the geometrical parameters in the region of the sulphur atom are the only ones to be influenced by the addition of the d-orbitals.

The main energetic effect of adding d-orbitals to the sulphur atom is a lowering of the total energy of the species, which is to be expected from a purely variational point of view, as the addition of extra basis functions allows more flexibility to the wave function description.

The largest percentage change in an orbital energy is that of orbital 23, by 2.3%. The d-orbital contribution is not of s-character in this case: the largest coefficient is that of the d_{xy}-function, with a slightly smaller d_{yy} contribution. The d_{xx}- and d_{zz}-functions have negligible coefficients. As there are also large exponents on the $2p_y$-functions on C2, it seems likely that the C-S bonding displays some (d-p) π type interaction.

The effect of the additional functions on the sulphur can also be noted in the charge distribution figures from a Mulliken population analysis. Considering the total atomic charges first, there is little or no difference in the charges on C1, O3, H4, O6, H7 and H9, but the difference in charge along the C2-S5 bond is reduced considerably, as the charges on C2 and S5 are reduced from -0.23 to -0.16 and 0.05 to -0.01 respectively. The overlap between C2 and S5 shows a slight increase in population, as does the sulphur atom occupancy. The number of electrons associated with C2 decreases. Thus, the basic trend in this region, upon addition of d-functions, is that fewer electrons are associated with C2 near its nucleus and more are associated with S5. The net effect is to reduce the negative charge on C2. This property of the sulphur atom may be thought of as helping to stabilise any anion formed by the removal of the proton, H7. This point will be discussed more fully later on.

A further calculation was made using the 6-31G** basis set , with the 3-21G* optimised geometry. This was made possible by an allocation of resources on the CYBER 205 at Manchester. Time limits prevented a complete geometry optimisation, but the improved basis set should shed further light on the electronic structure of the molecule, albeit at the 3-21G* geometry.

The addition of d-functions to the carbons and oxygens and p-functions to the hydrogens resulted in a lowering of the total energy of the species by 3.133 a.u. The wave function was similar to the 3-21G* one, showing significant participation of the d-functions of sulphur in some of the bonding orbitals. Orbital 23 of the wave function in this basis also shows a significant contribution from the d-function on the sulphur and the p-function of C2. The total amount

of electron participation in the d-functions on the sulphur for this basis set is 0.094 electrons as opposed to 0.127 electrons with the 3-21G* basis set. In fact, the total for C1, of 0.093, and C2, of 0.095, are similar to that of the sulphur.

The substrate model was further investigated by the use of two substitutions: the replacement of the thionyl group by a hydroxy group and by a proton, creating two new molecules in addition to A. The optimised geometries of these three molecules differ markedly: upon replacing the sulphur, the carbon-carbon bond is shortened in both cases.

The main change in the substrate model upon replacing the thiol group with a hydroxy group, apart from the obvious change in charge on X5, is the alteration of the charge on carbon C2 from -0.16 to 0.34. The more electronegative oxygen has drained the carbon of some electronic charge, making it more positive. This can be seen in the overlap population of these two species: the population centred at C2 is 5.53 and 5.01 respectively for molecule A and the hydrogen-substituted molecule A. The rest of the molecule is comparable to the substituted version, except that the C2-O6 bond overlap has changed from 0.42 to 0.47.

The hydrogen-substituted compound has very similar charges on the atoms to the thionyl compound, except that the protons H4 and H7 are slightly less positive. H7 is also less positive in the hydroxy-substituted molecule. If we look at the overlap population in this case, we find that there is a difference in the overlap on C2, as before, which is reduced from 5.53 to 5.31. Also the C2-O6 bond overlap has increased from 0.42 to 0.50.

The bond overlap of H7 to C2 is constant for all three molecules, being 0.70, and the overlap centred on H7 itself rises from 0.51 for molecule A to 0.53 for the hydrogen substituent and then to 0.58 for the hydroxy molecule. Thus, it does not seem, (especially if we also consider the bond distance data), as if the enzymic reaction gains any advantage from a weakening of the aldehydic proton binding - if anything this is strengthened. However the energetic stabilisation of molecule C with reference to molecule A seems a justification for the use of the hemithioacetal as an initial substrate. This characteristic does indeed seem to be a relative stabilisation of C, rather than a destabilisation of A, as the aldehydic proton does seem to be more securely bound to C2 in the sulphur-containing molecule A than in the other two cases, so the removal of the proton would be more difficult. This, then, is compensated for by the increased stability of the created anion, molecule C.

It appears that the d-functions do allow the molecule to indulge in some d-p interaction, but that it is not as important in defining the chemistry of this compound as was originally thought. It would be interesting to attempt a geometry optimisation of this molecule (and others in this model scheme) at the 6-31G** level, as recommended by Wolfe et al. [158], to see if this made much difference to the occupancy of the orbitals, or to the geometry of the molecule.

6.4 The Anionic Species

There are five anionic molecules in this model scheme: four with a single, overall, negative charge and one dianion:

The first two species, molecules B and C, are obtained by removing a proton from the substrate analogue compound, A. Molecule B results from the removal of the hydroxy proton, keeping the sp^3 carbon, C2; molecule C, on the other hand is a planar molecule, with two sp^2 carbons. The formal structures pictured in Figure 6-1 imply that the excess negative charge in these molecules will be located mainly on the oxygen atoms, but on different oxygens.

The third anion is the dianionic species, D. This can be obtained from either B or C, by the removal of the relevant hydrogen atom. The negative charges in this case are most likely to be found on the oxygen atoms, with a carbon-carbon double bonded, planar structure to the molecule.

The last two molecules, F and G, are singly charged species again and, like B and C, there is one planar species, molecule F, and one with an sp^3 carbon, molecule G. These structures are also likely to have the excess charge located on different oxygens. Molecule F can be the result of either a proton abstraction from the enediol intermediate, E, or of addition of a proton to molecule D. Molecule G, on the other hand can be the resultant only of an addition to the dianion. The product analogue, H, is than derived from either of these by a proton addition to the appropriate atom.

The anions B, D and G proved to be unoptimisable with the 3-21G* basis set: they broke up into fragments in which the SH^- structure is eschewed from the carbon atom.

The remaining parts after the "geometry optimisation" in the case of molecule B were cis-glyoxal and SH^-, but an STO-3G optimisation proved to be more stable. The resultant geometry was used as a starting point for a further attempt at a 3-21G* calculation, but this again proved to be unsuccessful. The STO-3G-optimised structure is listed in Table 6-3 and shown in Figure 6-3. Even in this state, the C-S bond length is very long, at 2.01 A. This may explain the instability when using the better basis set, in which the sulphur atom is not described by such a primitive set of functions and can accept more electronic charge, resulting in the cission of this bond.

Molecule G also has a stationary point on the STO-3G surface, but again, this could not be found using the 3-21G and the 3-21G* basis sets. The STO-3G geometry is listed in Table 6-4 and shown in Figure 6-4. The fragmented "products" of the optimisation procedure were formaldehyde, carbon monoxide and the thiol anion again.

A similar state of affairs holds for molecule D, the double anionic species. This gave an optimised geometry result at the STO-3G level, presented in Table 6-5 and Figure 6-5, but no stationary point was evident using the 3-21G and 3-21G* basis sets. Again, the C-S bond length is very long, being 2.10 A in the STO-3G structure. This weak bond, along with the additional basis functions available to the higher level calculations seem to combine, as before, to make the thionyl anion and other fragments more stable than the complete molecule, so the S-C bond is sundered by the optimising calculations.

The actual STO-3G wave functions for these three molecules are very poor indeed: molecules B and G have only 22 bound molecular orbitals and molecule D has 15 bound orbitals, whereas all of them should have 24 occupied, bound orbitals. Therefore, there is not much information which can be said to be reliably obtainable from these wave functions.

From the above calculations, it seems safe to assume that structures such as B, D and G do not play a role in the glyoxalase I reaction mechanism. It could be argued that stabilisation by a cation may occur in the enzyme, though; for instance the zinc atom of the glyoxalase I active centre could be chelated by the two oxygens. Lavery and Pullman have investigated the model reaction scheme in conjunction with a magnesium ion, finding that this ion reduced the large energy barrier between certain of the species. A similar scheme was tried with the reduced model system used in this work. Unfortunately, the constraints of computer disk space availability rendered this study virtually impossible with the 3-21G* basis set on the St Andrews VAX. Therefore a preliminary investigation, using the STO-3G basis set was undertaken. This will be discussed, along with details of the other molecules plus magnesium, later in this chapter.

The first stable anionic species, molecule C in the Lavery and Pullman scheme, has been studied with the 3-21G, 3-21G* and 6-31G** basis sets. As before, in the substrate analogue case, the 6-31G** calculation used the 3-21G* geometry. The results of the full geometry optimisations with the two smaller basis sets are listed in Table 6-6 and the 3-21G* basis set geometry is drawn in Figure 6-6.

Molecule C is a planar molecule, formed by the removal of the proton from the C2 carbon of the substrate analogue. The reaction then must proceed via the enediol intermediate, assuming that the dianion is unstable. This is energetically more favourable, also, as the enediol intermediate lies lower in energy than C, whereas the dianion has a higher energy than C on the STO-3G surface, which is the only one studied which seems to have anything resembling a stationary point on it. The addition of a proton to this anionic species releases almost as much energy as was required to form it from the substrate, as the enediol is of comparable energy to the substrate and product models.

Molecule C is formed by the removal of the proton H7 from molecule A. Upon removing this proton, apart from the obvious change in the overall charge on the molecule, the C1-C2 bond length is shortened considerably and the C1-O3 bond lengthens. The conjugation evident in molecule C along the O3-C1-C2-S5 plane is responsible for these changes. The oxygen atoms hold most of the negative charge on molecule A, as would be expected due to their electronegativity, but C2 also has some negative charge. This charge on the carbon is reduced upon removal of the proton and formation of the planar anion. The negative charge on the sulphur increases.

The main difference between the optimised geometries of molecule C using the 3-21G basis set and using the 3-21G* basis set is the length of the C-S bond, as with molecule A. This bond length is shortened by 4%, suggesting that the d-orbitals are again strengthening this bond. The only other changes leading to a greater than 1% change are the shortening of the S-H bond length by 1.9% and the reduction of the C1-C2-O6 angle by 1.3%.

On examination of the energies of the molecular orbitals it is found that the largest difference in energy between the two wave functions is an increase of 0.106 a.u. in molecular orbital 1, due to the s-character contribution of the d-orbitals to the inner shell sulphur 1s-orbital. The change in energies of these orbitals is much more pronounced than the case of the substrate model: there are eight molecular orbitals with a percentage change in energy greater than 1%, with by far the greatest change occurring in the twenty fourth, and highest-occupied orbital itself, which is reduced in energy by 15%. All but two of the top 12 orbitals are reduced in energy relative to their 3-21G equivalents.

The highest occupied molecular orbital is composed solely of p_z components in the 3-21G case, with the addition of $3d_{xz}$ and $3d_{yz}$ components on the sulphur for the 3-21G* case. (The z-direction is perpendicular to the plane of the molecule). The highest occupied molecular orbital consists of isolated p-orbitals on the oxygens and the sulphur, with a π-bonding function of the opposite polarity on the two carbon atoms. The d-functions on the sulphur have the opposite sign to the 3p-functions and so can combine with the C2 $2p_z$ to give some (d-p) π bonding. There are also lower energy orbitals which have an in-phase d-p contribution on these two atoms.

The Mulliken population overlap for C2 and S5 is 0.25 and 0.29 for 3-21G and 3-21G* respectively. Also, the sulphur-sulphur overlap has increased from 15.66 to 15.74 and the carbon population has decreased slightly, from 6.03 to 5.90. This supports the idea that there is some (d-p) π bonding, as this type of bond is a dative one, in which the carbon p-orbitals donate to the empty sulphur d-orbitals.

The acidity of the aldehydic proton adjacent to the sulphur has been investigated by comparing the stability of the anion and the substrate for the various substituted compounds.

These substituted derivatives of molecule C do show some differences in their structure and populations. The bonding distances for the original molecule are somewhere between those of the hydroxy-substituted version and the hydrogen-substituted one in all cases. The addition of the thiol group, in place of hydrogen, contracts the C1-C2, C1-H4 and C2-O6 distances and widens the C1-O3 and O6-H8 lengths. Upon subsequent replacement of the thiol group by hydroxyl, there is a marked enhancement of all these effects, especially the carbon-oxygen distances.

The pattern of charges derived from the Mulliken population analysis is quite similar in these three molecules, (ignoring X5), except for the carbon, C2, which is similar in charge for molecule C and its hydrogen-substituted variant, with -0.11 and -0.14, but has a charge of 0.42 in the hydroxy-derivative, making it the most positive atom present, with an even greater charge than is on the hydroxyl proton H9 (0.39).

Molecule C has the highest dipole moment (4.26 D) of the model reaction scheme; this is also true for the hydroxy-substituted species (4.15 D). However, the hydrogen-substituted molecule has the lowest dipole in its group (1.85 D). The high dipole moment of this anion could be used to re-orientate the substituent after the removal of the proton, facilitating the following re-protonation at the other carbon atom.

The 3-21G* optimised geometry was used in a 6-31G** single point calculation. The reduction in energy when this increase in the flexibility of the basis is made, is 3.127 a.u. The sulphur d-function population is the greatest, as might be expected, with 0.10 electrons, then the carbon d-function populations on C1 and C2 are 0.087 and 0.079 respectively. the oxygen populations are just 0.022 and 0.038 respectively. The main contribution of the sulphur d-functions are the d_{xz} and d_{yz} functions in the π -character orbitals.

The other anionic species, molecule F, is an isomer of molecule C in which the proton of the hydroxyl has moved from O6 to O3. This can be achieved either via the Lavery and Pullman scheme, by a proton being added to molecule C to form the enediol species, followed by a proton abstraction from the other oxygen, or by a direct proton transfer between the two oxygens. The second mechanism can only be of importance if there is no steric hindrance, as there is in the case of a chelated metal ion. Here, the hydrogen of the hydroxyl would be pointing away from the other oxygen and would be unable to hop across to it. A transition state structure has been found for the transfer of this proton and will be discussed after the results of the study of molecule F are reported.

The geometry of this species has been fully optimised using the 3-21G and 3-21G* basis sets. The results are reported in Table 6-7 and the 3-21G* geometry is shown in Figure 6-7. The C-C bond length does not alter much in comparison with the isomeric molecule C, but there is a noticeable shortening of the bond between carbon and the negative oxygen. This oxygen was attached to C1 in molecule C, but, as the proton has transferred to it to from F, the other oxygen, O6, has become the negative one. The adjacent sulphur strengthens the

interaction between C2 and O6 compared to that of C1 and O3 in molecule C. This is illustrated by examining the Mulliken population total atomic charges for the two molecules. In molecule C the carbon attached to the sulphur, C2, is also attached to a hydroxyl oxygen and has an overall charge of -0.11, similar to the -0.10 found on the sulphur atom. The carbon attached to the other oxygen has an overall charge of 0.19. In contrast to this, C2 in molecule F has a charge of 0.27, significantly larger than that of C1 in molecule C and the sulphur has a charge of -0.28. This is the largest negative charge found on the sulphur in this reaction scheme. C1 has the same value of -0.11 as C2 in molecule C and the negative oxygen, O6 has a similar charge to that of O3 in molecule C, with -0.79, as opposed to -0.81.

If one examines the all-electron overlap populations of the molecules, it is obvious that the sulphur atom is responsible for the increased positive charge on C2, relative to C1 of molecule C, as its own population increases from 15.74 to 16.07 and the C2-S5 bond population also increases from 0.29 to 0.33. Again, the d_{xz}- and the d_{yz}-functions play a part in the conjugated π-bonding from O3 and O6, through C1 and C2 to the sulphur. The other d-functions seem to contribute to the sulphur s-functions. The overall effect of the d-functions is less than for molecule C, in which the sulphur is next to the negative carbon, C2: the overall population of the d-functions is 0.138 electrons in molecule C and 0.108 in molecule F.

Thus, the net consequence of transferring the proton from O6 to O3 is to reverse the charge on the two carbon atoms, using the sulphur atom to make C2 more positive than C1 previously was. This phenomenon may be used by the enzyme to donate the second proton - which was originally attached to C2 - to C1 to complete the isomerisation reaction.

A comparison of the 3-21G and 3-21G* optimised geometry results shows that molecule F is affected much more by the d-functions' presence on the sulphur atom than any other species in the scheme. The C2-S5 bond length is reduced by 12% and the S5-H7 bond length by 2.2%, compared with about 4% and 1.9% for the other intermediates. The C2-O6 bond length is extended by 3.2% also, whereas the other molecules' C2-O6 bond lengths do not change by more than 1%. There are also large changes in the bond angles C1-C2-S5 and C1-C2-O6 which increase by 6.9% and decrease by 7.0% respectively. The net effect of this bond angle movement is that the O6 oxygen moves away from the O3 oxygen and the S5 sulphur moves closer to the H4 atom and further away from O6, as the S5-C2-O6 bond angle also increases by 2.2%. The 3-21G* optimised structure in the vicinity of C2 is a more even arrangement, with angles closer to 120° than the distorted 3-21G geometry. The sulphur atom in particular is bent away from the C1-C2 axis, as the C1-C2-S5 bond angle is only 105.8°.

The total d-orbital population of molecule F is only 0.10821, which is the smallest total in the whole model scheme, notwithstanding its large effect on the optimised geometry. However, it does have the largest totals of s- and p-function populations on the sulphur of any of the intermediates, reflecting the fact that it has the most negative sulphur atom in the scheme, according to the Mulliken analysis.

There is considerable difference in the orbital energies of the 3-21G and the 3-21G* results, compared to the other molecules, in which the greatest percentage difference was of the order of 2% to 4%. All the top half of the occupied molecular orbitals, from orbitals 13 to 24, have a greater than 1% change in energy. The highest occupied molecular orbital has the largest change, of 31%, and yet the change in

the total energy is the smallest found in the reaction scheme upon going from the 3-21G to the 3-21G* geometries. This can be explained as follows: the large change in the molecular geometry of molecule F upon the addition of the d-functions to the basis prompts a large change in the nuclear repulsion energy term, of 6.89 a.u., which is balanced by an even larger change in the electronic energy, of -7.00 a.u., giving a small total energy reduction compared to the other molecules. These all have changes of around 2 a.u. in the nuclear repulsion energy term and a reduction of greater magnitude in the electronic energy. Thus, the addition of the d-functions to the basis of molecule F has a profound effect on the resultant optimised geometry wave function.

Upon replacing the thiol group, S5-H7, with a hydroxy group or a hydrogen atom, the carbon-carbon and carbon-oxygen bonds expand and the other bond lengths are fairly constant (ignoring the ones including the sulphur atom). The Mulliken population results are more or less the same except at C2, where a charge of 0.27 for the thiol derivative is replaced by 0.19 for the hydrogen derivative and 0.63 for the hydroxy derivative, where it is adjacent to two electronegative oxygens. As might be expected, the overlap is least on the hydroxy derivative C2 (4.58). The other overlap populations - again ignoring the substituted group - are very similar: in fact, they are closer to the 3-21G* optimised geometry values of molecule F than the 3-21G optimised geometry values are. This is also true of the Mulliken total atomic charges, in which the 3-21G set are the furthest removed from the 3-21G* ones, except for the C2 populations.

A further single point calculation using the 6-31G** basis set at the 3-21G* optimised geometry resulted in a further energy reduction of 3.129 a.u. The wave function resembled the 3-21G* one, with contributions from the d_{xz}- and d_{yz}-functions in the conjugated π-bonding orbitals, consisting mainly of p_z-functions. The C2 atom also seems to have a considerable population in its d-functions. These seem to participate in the bonding description, rather than the lower energy core orbitals, especially the p_z-containing orbitals.

6.5 The Transition State Species: Molecule CF

Pauling's original idea [15], that an enzyme is bound most strongly to the transition state of the reaction it catalyses has enjoyed a resurgence of popularity since its revival by Wolfenden [160]. This author reviewed the subject thoroughly, listing all the then-known examples in which the inhibitors of the particular enzyme could be thought of as transition state analogues. This paper stimulated much work based on the idea, which is now firmly entrenched in enzyme chemistry.

One problem with many of the discussions of so-called transition states is that they focus not upon transition states per se, but on reactive intermediate species. A reactive intermediate is a transient species often, but not necessarily, of high energy, which exists on a reaction coordinate between the reactants and products. It differs from a true transition state in that it is a stationary point which is a minimum on the reaction hypersurface, whereas the transition state is a maximum in one coordinate direction, that is, it is a saddle point on

the surface.

The theory of transition state analogue inhibitors states that an effective inhibitor can be produced if it is similar in structure to that of the transition state of the enzyme-catalysed reaction [12]. If we consider the reaction scheme devised from the model compounds, we can see that the most effective inhibitors will probably have a structure not similar to the enediol intermediate, as suggested by Douglas and Nadvi [102], but to one of the anionic molecules C and F. These are high energy, transient species which an enzyme would have to bind efficiently if the mechanism proceeded via alternate deprotonation and protonation, as seems likely.

One true transition state has been found on the reaction hypersurface; this is the transition state for the transfer of a proton between the two oxygen atoms of molecules C and F. If glyoxalase I catalyses this process, rather than allowing the formation of the enediol species, E, then this species would be the most tightly bound one. Unfortunately, it is difficult to find a compound which has a similar structure to this transition state, having a proton intermediate between two oxygen atoms; it is not very surprising that the majority of good inhibitors of glyoxalase I contain a structure similar to the O-C-C-O skeleton of molecules C and F, or similar to the enediol species. By this argument one would expect planar α-hydroxyketone compounds to be better inhibitors than planar enediol ones. However, this is a little too simplistic, as the ease of deprotonation of a hydroxyl group to give an anionic species may also be important in the case of enediol-like compounds.

The transition state species in this scheme, molecule CF, was found by using the transition state search algorithm in the Berny optimisation program in GAUSSIAN 82. The search was carried out on the 3-21G* surface. The resulting energy of CF was -621.5216 a.u., compared with -621.5358 a.u., -621.5513 a.u., and -622.1378 a.u. for molecules C, F and E respectively. The resultant optimised geometry is listed in Table 6-8 and shown in Figure 6-8.

The geometry of molecule CF might be expected to be halfway between the geometries of C and F and, indeed, this is generally the case. Most of the bond lengths fall between the two optimised values for C and F, with one notable exception: the carbon-carbon bond length is shorter than the corresponding lengths for these molecules, being 1.328 A. This can be seen by looking at the Mulliken overlap populations of these species: the C1-C2 overlap is far greater in CF (0.80) than in either C (0.48) or F (0.53). The other obvious differences in atomic distances are the oxygen-hydrogen distances, which lengthen considerably, from 0.97 to 1.2 A

The part of the molecule which is to the opposite side of the C-C bond from the oxygens is comparable in structure to molecules C and F: The sulphur atom of CF is similar to that of molecule C - the bond distances involving this atom are almost the same as the respective lengths in molecule C and the bond angle C2-S5-H7 is very similar in these two species. The other end of the CF structure is nearer to molecule F in shape: the hydrogen placement involves a slightly larger C1-H4 bond, but the same C2-C1-H4 angle.

The two oxygens, on the other side of the molecule have both moved in towards the intermediate hydrogen; the C-C-O bond angles have both become much more acute, being around 112° instead of the previous values near to 120°. The C-O-H angles have also shortened from around 100° to around 80° and the O-H distances have lengthened, as mentioned before. The net effect of the changes in the C-C, C-O and O-H geometry is to place the hydrogen fairly centrally, stabilised by two weak bonding interactions. It is slightly closer to O6 than to O3, this marginally stronger association being reflected by the O-H Mulliken overlap populations, of which the O6-H8 is the larger.

The d-orbital population of this species is 0.136, which is less than that for molecule C, but more than molecule F. The C2 overlap population is 5.44, which is also midway between C and F, as is the atomic charge on C2 of 0.01.

Perhaps the most interesting feature of this species is the position of the hydrogen H8, which is not bonded strongly to any particular atom, but has weak bonding association with both O3 and O6, with overlap populations of 0.31 and 0.32. This compares with 0.51 and 0.53 for the hydroxyl bonds of the other two molecules.

The alternative route from C to F involves the enediol, E. This is a very different species from the CF transition state, being a stable intermediate.

6.6 The Enediol Model: Molecule E

The enediol species, molecule E, has been proposed as a target species for the design of "transition state" analogue inhibitors of glyoxalase I. It is a planar molecule in which the most stable arrangement of the hydroxy groups is the conformation in which the hydrogen atoms point away from the sulphur atom. The energy of this molecule is -622.1378 a.u., which is comparable to the -622.1470 a.u. obtained for the substrate analogue, molecule A, and below the energies of the anionic species C and F (-621.5358 and -621.5513 a.u. respectively). The optimised geometry, reported in Table 6-9, is for the 3-21G and 3-21G* basis sets. Figure 6-9 depicts the optimised geometry from the 3-21G* calculation.

An interesting feature of the structure of this intermediate is the large included angle of the C1 hydroxyl group, C1-O3-H9. This is 112.7°, whereas the hydroxyl group angles in the 3-21G basis set are generally around 109°.

The addition of a proton to the O3 oxygen of molecule C results in the formation of this enediol species. This is a favourable reaction liberating 0.602 a.u. The electrostatic potential map in the plane of molecule C has its most negative region in the neighbourhood of the subsequent hydroxyl proton's position.

The change in geometry associated with this reaction demonstrates the transformation of a high energy metastable species into a stable intermediate. There is no ambiguity about the structure: the carbon-carbon distance is 1.312 A, a definite unsaturated bond and the carbon-oxygen bonds are both long, being around 1.4 A, as opposed to

molecule C, where one of these was 1.3 A (C1-O3) and the other was 1.43 A (C2-O6). The carbon-sulphur distance also shortens, as the sulphur participates more in the conjugated bonding at the expense of the oxygen atoms.

The addition of d-functions to the sulphur, to convert the 3-21G basis set representation to a 3-21G* one results in a shortening of the C2-S5 bond by 3% and a shortening of the S5-H7 one by 1.8%. All the other changes are of less than 1%. This intermediate has the greatest total population of d-orbitals of any molecule in the reaction scheme, and has the shortest C2-S5 bond, with the greatest overlap population, of 0.40.

The change in energy of the individual molecular orbitals can be divided into two groups as before: there are changes brought about by an addition of s-character to various orbitals, particularly the first orbital, which is predominantly sulphur 1s, due to equal contributions from the xx, yy and zz functions; there are also changes brought about by an improved description of the π-bonding in the molecule. This latter effect is found in the higher occupied orbitals, with its largest percentage change in the twenty-second and twenty-fourth orbitals. Molecular orbital 24, the highest occupied molecular orbital consists of just p_z contributions of roughly the same magnitude on C1, C2, O3, O6 and S5 in the 3-21G wave function. (The z-axis is perpendicular to the plane of the molecule). The 3-21G* twenty-fourth molecular orbital is similar, except that it also contains a contribution from the d_{xz} and d_{yz}-functions. The orbitals which can be classed as π-bonding states, with p_z contributions are found to all have d_{xz} and d_{yz} contributions also.

The carbon-carbon bond of the original enediol species is larger by a small amount than that of either the hydroxy- or the hydrogen-substituted species. The other bonds are generally comparable to the hydrogen-replaced values, except for the positioning of the H8 atom, which is the hydroxyl proton pointing towards the other cis hydroxyl group. This is nearer to the adjacent hydroxyl oxygen, O3, than in the other two cases, with a tighter C2-O6-H8 angle and a larger O6-H8 bond length. The other hydroxyl internal angle, C1-O3-H9 is decidedly large in all three of the enediol molecules, being around 112.5°, The C1-C2-O6 angle is also tighter than for the two derivatives, pushing the H8 proton even closer to O3, but all three angles are wider than a normal sp^2 carbon internal angle, due to the strain of the eclipsed conformation. This compares with the much more acute values found for the substrate models, which are of course sp^3 carbon centres. They are also wider than the corresponding ones for the molecule C derivatives, where the C1-C2-O6 angles are about 117°.

A further calculation on the enediol was run using the 3-21G* geometry and the 6-31G** basis set. The energy was reduced to -625.2581 a.u., a reduction of 3.120 a.u. The form of the wave function was similar to the 3-21G* one, with d_{xz} and d_{yz} contributions to the π-bonding orbitals. The d-functions placed on carbon C2 were almost as well populated as those on S5, having 0.092 electrons as opposed to 0.10 electrons. The majority of this contribution was to the p_z π-bonding orbitals as d_{xz} and d_{yz} contributions as noted for the other molecules.

6.7 The Product Analogue: Molecule H

The final product of the model reaction scheme is molecule H. This has an sp^3 carbon at C1 and an sp^2 one at C2 - the opposite to the original molecule A. Unlike molecule A, it is a symmetrical molecule, of Cs symmetry. It has been optimised fully at both the 3-21G and 3-21G* levels, the results of that optimisations are presented in Table 6-10 and the 3-21G* geometry is shown in Figure 6-10.

The carbon-carbon bond length is larger than that of molecule A, being 1.521 A, compared to 1.509 A. The carbonyl carbon-oxygen bond length, C2-O6, is slightly larger than the original carbonyl length for C1-O3 and the carbon-oxygen distance for the hydroxy group is also larger than for molecule A. The carbon-sulphur is almost as short as the enediol carbon-sulphur distance, being 1.766 A and the total d-function contribution to the wave function is also second only to molecule E, being 0.140 electrons in total.

Molecule H is energetically more stable than the isomeric molecules A and E, in fact, it is the most stable molecule in the model scheme, the total energy of the species being -622.1608 a.u.

The main difference in the geometries of the 3-21G optimised and the 3-21G* optimised results is again the C2-S5 bond length, which is reduced by 4.4%, and the S5-H8 distance which is reduced by 1.9%. The sulphur atom and the carbonyl oxygen combine to pull electrons away from C2, making it quite positive, with a Mulliken atomic charge of 0.30. In fact, this is the most positive C2 carbon in the whole scheme, just as C2 is the most negative for molecule A. This reversal of charge on carbon when comparing the substrate and product models is

also noticeable on C1, where a charge of 0.31 becomes one of -0.14 as C1 changes from an sp^2 carbon to an sp^3 one. This is an enhancement of the effect observed for the conversion of the isomers C to F, in which there is a similar reversal of charge at these centres.

The highest occupied molecular orbital shows a contribution from the d_{xz}- and d_{yz}-functions, as do all the other orbitals with contributions from the p_z-functions. The greatest change in energy, upon adding the d-functions to the sulphur atom is a 1.5% reduction in the energy of molecular orbital 23, the second highest occupied orbital. This has a contribution from the d_{xy}-function, but its major contributions are p_x coefficients on S5 and O6. Molecular orbital 22 increases in energy by 1.4% and has contributions from sulphur d_{xz}- and d_{yz}- as well as from the p_z-functions on the other atoms.

The total population of the d-functions is 0.140 for molecule H, second only to molecule E, the enediol species. Both these molecules are neutral and have a planar S-C-O region in their structure. The interaction between the sp^2 carbon and the sulphur d-orbitals is strong in these molecules; the C2-S5 overlap population is large for these species too.

The effect of the addition of d-orbitals is also apparent in a comparison of the Mulliken total atomic charges. The charges on C1, O3, H4 and H7 are about the same for both 3-21G and 3-21G* wave functions, but the charge on C2 and S5 alters: that on C2 increases from 0.20 to 0.30 and the charge on S5 decreases from 0.10 to 0.02, as more electrons occupy the sulphur orbitals at the expense of the carbon. The C2-S5 overlap population increases from 0.31 to 0.38 and the population on C2 is reduced (5.44 -> 5.24) and that on S5 is increased (15.61 -> 15.75).

If the thiol group is replaced by a hydroxy group or a hydrogen atom, the resulting change in geometry and electron distribution is evident in that region of the molecule only: the hydroxy side of the molecule is unaffected. The C1-C2 carbon-carbon bond shrinks considerably when the sulphur is replaced b an oxygen atom. Likewise, the total atomic charge pattern is only changed in the region of C2, (ignoring the X5 side group). C2 has a charge of 0.30 for the thiol and hydrogen derivatives, but a charge of 0.80 for the hydroxy species. The overlap population on C2 is also decreased in the hydroxy derivative, from 5.07 to 4.54.

The use of the 6-31G** basis set at the 3-21G* optimised geometry results in a reduction of the energy by 3.248 a.u., to -625.2948 a.u. The largest population of d-functions are again found on C2 and S5, being 0.11 and 0.10 electrons respectively.

6.8 Electrostatic Potential Calculations

The molecular electrostatic potential maps in the plane of the O-C-C-O framework were calculated for the optimised geometries of the model compounds, using a version of the DENPOT program [139], incorporated into the GAUSSIAN 80 program [138]. The van der Waals surface electrostatic potentials were also calculated, using another version of GAUSSIAN 80 [143]. These were plotted using the colour plotting program described in Chapter 5. The resulting pictures are shown in Figures 6-11 to 6-18.

The main diagnostic feature of the planar diagrams is the position of the negative potential wells, which can be equated to the oxygen and sulphur lone pairs. At the start of the model enzyme reaction, molecule A has a large well around the O3 carbonyl oxygen and a smaller one on the excluded side of C2-O6-H8 at a distance of about 4.8 A. Upon removal of the H7 proton, to form the anionic species, molecule C, there is an extensive change in the structure of the MEP. As is the case for most anionic species, the surrounding electrostatic potential is negative in almost every direction, as the overall negative charge will attract a positive point charge from any side. However, the minima are again associated strongly with the two oxygens, as one would expect and are about the same distance apart, at 4.7 A. The shape of the wells has changed, though. Molecule A has a substantial ridge of positive potential between the two oxygens, due to the H9 hydrogen, but this is very much reduced in molecule C so that the positive potential does not protrude even as far as the van der Waals surface of this atom.

If we now consider molecule E, the enediol species, we find a different pattern again: the minima associated with the oxygens have shrunk to a quite small size compared to molecule A. Not only has the addition of a proton to the carbonyl oxygen reduced the size of this atom's negative region, but also the hydroxy oxygen, O6, has a reduced negative region. This is because the atom H7 is now in the plane of the molecule, pinching the negative region between its repulsive, positive contribution to the MEP and that of the H8 hydrogen. The distance between the negative regions for this species is about 3.4 A, as the position of the minimum associated with O3 has changed because of its attendant hydrogen atom, H9.

Molecule F, being another anionic species does not have any appreciable positive potential outside the van der Waals surface in the plane of the molecule. Again, the deepest part of the negative areas are about 4.5 A apart and the shape of the map is reminiscent of that of the isomeric molecule C, except that the regions associated with the oxygens are reversed.

The final species, molecule H, has the largest negative region of any neutral species, associated with its hydroxyl oxygen and a more compact carbonyl region of a similar size to the original hydroxy oxygen region of molecule A. The negative wells are again about 4.3 A apart.

The electrostatic potential on the van der Waals Connolly surface has also been calculated for certain members of the reaction scheme. The calculation was restricted to the neutral members of the scheme, as the anionic species cannot be directly compared to the neutral inhibitor molecules. The results are shown in Photographs 1 to 4.

The van der Waals surface electrostatic potential plots have negative regions associated with the oxygen and sulphur atoms, as one would expect. The enediol species, molecule E, also has a negative region between the two carbon atoms, whereas the substrate and product analogues, molecules A and H, have a positive potential in this area. Thus the conversion of the bond between the carbons of the substrate to an unsaturated, rigid bond is accompanied by a complete change in the potential of this area. A "transition state" inhibitor of this enzyme might be expected to have a similar potential in this region, or at least it should not have the positive regions. This will be discussed in the next chapter.

6.9 Binding Studies With Magnesium

The UMRCC CYBER 205 has a version of the GAMESS program which includes the MINIT optimisation routine developed by Bell and Crighton [161]. This has proved to be a better algorithm than the Berny one for optimising the geometry of more unusual systems [162]. The geometries of the molecules in this reaction scheme have been optimised with the addition of a magnesium dication between the two oxygen atoms, as an attempt to mimic the glyoxalase I active site zinc atom. It has been shown by Mannervik's group that the magnesium-substituted enzyme works just as well as the original zinc version [64]. Lavery and Pullman [149] also studied the effect of magnesium on the model compounds in their partially-optimised scheme. They found that the addition of magnesium resulted in a large drop in the activation energy needed to remove the protons, stabilising the anionic species relative to the neutral ones.

The energies obtained from the 3-21G* optimisation of these molecules are reported in Table 6-11, and the optimised geometries are listed in Tables 6-12 to 6-19 and shown in Figures 6-11 to 6-18.

As can be seen from Table 6-11, this work seems to support that of Lavery and Pullman in that the activation energies of all the steps are reduced considerably, as the binding energy of the anions with magnesium is greater than the binding energy for the neutral species. (There are no binding energy results for the molecules B, D and G, as these were unstable without the magnesium ion). The binding energies with magnesium are -0.67 a.u. for the anionic species C and F, -0.28 a.u. for A and E and slightly more, -0.30 a.u., for the product analogue, H.

The actual positioning of the magnesium ion in these full optimisation results, compared to the partially optimised work of Lavery and Pullman, is more central, with shorter bonding distances of about 1.8 A for Mg^{2+}-O^{-} interactions and about 1.9 A for magnesium to carbonyl or hydroxyl oxygens, as opposed to 1.77, 1.97 and 2.00 A.

The most favoured pathway, energetically, is still via C, E and F as found for the original scheme, but the route via D is possible this time, as the deprotonation reaction requires 0.366 a.u. and the protonation reaction of imidazole, which is probably part of the base of the active site (histidine) [69], is -0.397 a.u. in the 3-21G basis. However, it seems that it would be difficult to effect this second deprotonation if there was any proton donating species present, as the C->E protonation reaction energy is -0.0204 a.u., a much more favourable route. For instance, the deprotonation of H_3O^{+} in this basis gives an energy of 0.305 a.u. The deprotonation reaction of E to give F requires only 0.189 a.u. now, compared with 0.587 a.u. for the original scheme and could easily be accomplished by a base such as histidine.

There is one important qualification which must be placed on the above scheme. Not one of the magnesium-associated molecules has a good wave function in that there are always a number of bound, unoccupied molecular orbitals. Molecule D, the dianion, has only one extra bound orbital, but the anionic species have 4 or 5 and the neutral (previous to addition of Mg^{2+}) species A and H have 13 and 12 bound unoccupied orbitals! Therefore it is unfortunately the case that little emphasis can be placed upon this scheme as being of interpretive value.

6.10 Addition Of Alkyl Groups To Sulphur

As the amount of computer time available increased, it became possible to expand the model scheme to include calculations on larger species, as well as the magnesium-associated species. The stable intermediates from the original scheme were used, replacing the thiol hydrogen with methyl, or, in the case of the substrate and product models, with ethyl groups. The thiol group is therefore converted into a thioether group. The resultant optimised geometries, using the 3-21G* basis set are listed in Tables 6-20 to 6-24. An attempt was also made to optimise the molecule B structure with an ethyl group replacing the hydrogen, but this proved to be unstable, throwing out an SEt^- species. It is immediately apparent, from inspection of these values, that the replacement of the hydrogen with larger alkyl groups does not affect the geometry of the molecule in any great manner: the changes in bond lengths are very small, being less than 1%. The charges and overlap populations from the Mulliken analysis are very similar to the original molecules' values, except that the charge on the sulphur is generally increased by about 0.25 electrostatic units, due to the inductive effect of the adjacent alkyl group in comparison with a single hydrogen. The main point is that the O-C-C-O part of the molecule is unaffected by the presence or absence of the alkyl group, vindicating the original decision to replace the glutathione region with a thiol group.

Table 6-1 Total Energies And Energy Differences Of Molecules In The Reaction Scheme, In Atomic Units.

	3-21G	3-21G*	SH->OH	SH->H
A	-622.0374	-622.1470	-300.9307	-226.4844
C	-621.4236	-621.5358	-300.2996	-225.8609
E	-622.0237	-622.1378	-300.9212	-226.4747
F	-621.4505	-621.5513	-300.3331	-225.8609
H	-622.0472	-622.1608	-300.9634	-226.4844
A->C	+0.6138	+0.6112	+0.6311	+0.6235
C->E	-0.6001	-0.6020	-0.6216	-0.6138
E->F	+0.5732	+0.5865	+0.5881	+0.6138
F->H	-0.5967	-0.6095	-0.6303	-0.6235
A->E	+0.0137	+0.0092	+0.0095	+0.0097
E->H	-0.0135	-0.0230	-0.0422	-0.0097
C->F	-0.0269	-0.0155	-0.0335	0.0000
A->H	-0.0098	-0.0138	-0.0327	0.0000

Table 6-2 Optimised Geometries Of Molecule A And Some Derivatives, Replacing Thionyl, In Angstroms And Degrees.

	3-21G	3-21G*	SH->OH	SH->H
C1-C2	1.509	1.517	1.515	1.509
C1-O3	1.208	1.208	1.208	1.210
C1-H4	1.080	1.081	1.078	1.083
C2-X5	1.906	1.828	1.419	1.085
C2-O6	1.404	1.416	1.401	1.426
C2-H7	1.078	1.081	1.082	1.085
X5-H8	1.350	1.325	0.967	-----
O3-H9	2.088	2.074	2.086	2.097
O6-H9	0.972	0.972	0.971	0.970
C2-C1-O3	120.5	120.8	121.1	121.6
C2-C1-H4	116.2	116.2	115.3	116.1
O3-C1-H4	123.3	123.0	123.6	122.3
C1-C2-X5	105.8	107.5	102.9	108.6
C1-C2-O6	111.1	110.0	110.2	110.4
X5-C2-O6	111.0	111.6	113.5	110.7
C1-C2-H7	110.8	109.0	110.2	108.7
X5-C2-H7	106.6	108.4	110.7	107.7
O6-C2-H7	111.4	110.2	109.2	110.7
C2-X5-H8	95.5	95.2	111.1	-----
C2-O6-H9	109.3	109.1	109.9	108.5
O3-C1-C2-X5	120.5	121.7	121.3	121.5
O3-C1-C2-H7	-124.4	-121.0	-120.7	-121.6

Table 6-3 Optimised Geometry Of Molecule B, Using STO-3G Basis, In Angstroms And Degrees.

	STO-3G
C1-C2	1.597
C1-O3	1.222
C1-H4	1.115
C2-S5	2.015
C2-O6	1.282
C2-H7	1.124
S5-H8	1.329
C2-C1-O3	127.4
C2-C1-H4	113.9
O3-C1-H4	118.7
C1-C2-S5	97.3
C1-C2-O6	119.4
S5-C2-O6	115.8
C1-C2-H7	102.7
S5-C2-H7	96.9
O6-C2-H7	120.2
C2-S5-H8	95.4
O3-C1-C2-S5	125.2
O3-C1-C2-H7	-136.0
C1-C2-S5-H8	271.3

Table 6-4 Optimised Geometry Of Molecule G, Using STO-3G Basis, In Angstroms And Degrees.

	STO-3G
C1-C2	1.688
C1-O3	1.312
C1-H4	1.129
C2-S5	1.857
C2-O6	1.214
C1-H7	1.129
S5-H8	1.332
C2-C1-O3	116.9
C2-C1-H4	97.7
O3-C1-H4	119.0
C1-C2-S5	114.0
C1-C2-O6	128.9
S5-C2-O6	117.1
C2-C1-H7	97.7
O3-C1-H7	119.0
H4-C1-H7	102.4
C2-S5-H8	96.1
O6-C2-C1-H4	128.2
O6-C2-C1-H7	-128.2

Table 6-5 Optimised Geometry Of Molecule D, Using STO-3G Basis, In Angstroms And Degrees.

	STO-3G
C1-C2	1.356
C1-O3	1.353
C1-H4	1.117
C2-S5	2.100
C2-O6	1.287
S5-H7	1.336
C2-C1-O3	129.5
C2-C1-H4	112.4
O3-C1-H4	118.1
C1-C2-S5	108.1
C1-C2-O6	140.2
S5-C2-O6	111.7
C2-S5-H7	97.5

Table 6-6 Optimised Geometries Of Molecule C And Some Derivatives, Replacing Thionyl, In Angstroms And Degrees.

	3-21G	3-21G*	SH->OH	SH->H
C1-C2	1.334	1.338	1.324	1.340
C1-O3	1.301	1.300	1.327	1.298
C1-H4	1.093	1.095	1.089	1.100
C2-X5	1.848	1.776	1.408	1.069
C2-O6	1.427	1.433	1.417	1.443
X5-H7	1.349	1.324	0.965	-----
O3-H8	2.111	2.083	1.992	2.145
O6-H8	0.978	0.978	0.983	0.974
C2-C1-O3	122.8	123.2	120.9	125.3
C2-C1-H4	116.3	116.1	116.9	114.6
O3-C1-H4	121.0	120.7	114.9	120.1
C1-C2-X5	128.6	129.1	130.0	127.8
C1-C2-O6	118.7	117.2	117.1	117.1
X5-C2-O6	112.7	113.8	116.6	115.1
C2-X5-H7	94.9	94.9	106.2	-----
C2-O6-H8	100.4	100.3	99.4	100.9

Table 6-7 Optimised Geometries Of Molecule F And Some Derivatives, Replacing Thionyl, In Angstroms And Degrees.

	3-21G	3-21G*	SH->OH	SH->H
C1-C2	1.324	1.334	1.338	1.340
C1-O3	1.442	1.441	1.448	1.443
C1-H4	1.060	1.065	1.063	1.069
C2-X5	2.173	1.913	1.425	1.100
C2-O6	1.226	1.265	1.274	1.298
X5-H7	1.355	1.325	0.967	-----
O3-H8	0.967	0.970	0.970	0.974
O6-H8	2.588	2.312	2.262	2.145
C2-C1-O3	119.5	118.0	117.5	117.1
C2-C1-H4	126.2	127.1	126.7	127.8
O3-C1-H4	114.3	114.9	115.8	115.1
C1-C2-X5	105.8	113.1	115.5	114.6
C1-C2-O6	140.1	130.3	128.7	125.3
X5-C2-O6	114.1	116.6	115.9	120.1
C2-X5-H7	92.6	92.5	104.1	-----
C1-O3-H8	107.0	103.9	103.1	100.9

Table 6-8 Optimised Geometries Of Molecule CF, The Transition State Between Molecules C And F, In Angstroms And Degrees.

	3-21G*
C1-C2	1.328
C1-O3	1.382
C1-H4	1.073
C2-S5	1.775
C2-O6	1.385
S5-H7	1.324
O3-H8	1.288
O6-H8	1.210
C2-C1-O3	111.3
C2-C1-H4	126.9
O3-C1-H4	121.9
C1-C2-S5	129.3
C1-C2-O6	112.0
S5-C2-O6	118.7
C2-S5-H7	94.5
C1-O3-H8	88.0
C2-O6-H8	88.7
O3-H8-O6	140.1

Table 6-9 Optimised Geometries Of Molecule E And Some Derivatives, Replacing Thionyl, In Angstroms And Degrees.

	3-21G	3-21G*	SH->OH	SH->H
C1-C2	1.307	1.312	1.309	1.310
C1-O3	1.406	1.407	1.417	1.404
C1-H4	1.066	1.067	1.064	1.068
C2-X5	1.823	1.763	1.358	1.066
C2-O6	1.374	1.379	1.365	1.379
X5-H7	1.348	1.323	0.966	-----
O3-H8	2.183	2.156	2.170	2.245
O6-H8	0.971	0.970	0.968	0.968
O3-H9	0.962	0.961	0.961	0.961
C2-C1-O3	116.1	116.1	115.1	117.5
C2-C1-H4	123.9	124.1	123.6	123.2
O3-C1-H4	120.0	119.8	121.3	119.3
C1-C2-X5	124.3	124.8	124.3	123.0
C1-C2-O6	123.0	121.9	123.0	123.5
X5-C2-O6	112.8	113.4	112.8	113.5
C2-X5-H7	94.6	94.1	110.6	-----
C2-O6-H8	109.4	109.3	110.4	109.8
C1-O3-H9	112.7	112.7	112.0	112.8

Table 6-10 Optimised Geometries Of Molecule H And Some Derivatives, Replacing Thionyl, In Angstroms And Degrees.

	3-21G	3-21G*	SH->OH	SH->H
C1-C2	1.512	1.521	1.504	1.509
C1-O3	1.425	1.424	1.424	1.426
C1-H4	1.083	1.083	1.082	1.083
C2-X5	1.847	1.766	1.345	1.085
C2-O6	1.200	1.209	1.204	1.210
C1-H7	1.083	1.083	1.082	1.083
X5-H8	1.350	1.325	0.969	-----
O3-H9	0.970	0.970	0.969	0.970
O6-H9	2.123	2.074	2.147	2.097
C2-C1-O3	109.9	109.6	110.0	110.4
C2-C1-H4	108.7	108.9	108.3	108.7
O3-C1-H4	110.6	110.5	111.0	110.7
C1-C2-X5	113.9	115.2	111.8	116.1
C1-C2-O6	123.0	121.0	124.3	121.6
X5-C2-O6	123.0	123.9	123.9	122.3
C2-C1-H7	108.7	108.9	108.3	108.7
O3-C1-H7	110.6	110.5	111.0	110.7
H4-C1-H7	108.4	108.4	108.1	107.7
C2-X5-H8	95.8	95.0	112.8	-----
C1-O3-H9	109.4	108.9	109.4	108.4
O4-C2-C1-H7	121.1	121.0	121.5	121.6
O4-C2-C1-H8	-121.1	-121.0	-121.5	-121.6

Table 6-11 Total Energies And Energy Differences Of Molecules In The Reaction Scheme Plus Mg^{2+}, In Atomic Units.

	3-21G*
A	-820.1338
B	-819.9308
C	-819.9152
D	-819.5492
E	-820.1195
F	-819.9303
G	-819.9520
H	-820.1668
A->B	+0.2030
A->C	+0.2186
B->D	+0.3816
C->D	+0.3660
C->E	-0.2043
D->F	-0.3811
D->G	-0.4028
E->F	+0.1892
F->H	-0.2365
G->H	-0.2148
A->E	+0.0143
E->H	-0.0473
C->F	-0.0151
A->H	-0.0330

Table 6-12 Optimised Geometry Of Molecule A Plus Mg^{2+},

In Angstroms And Degrees.

	3-21G*
C1-C2	1.511
C1-O3	1.238
C1-H4	1.077
C2-S5	1.806
C2-O6	1.517
C2-H7	1.083
S5-H8	1.327
O6-H9	0.974
O3-Mg10	1.908
O6-Mg10	1.922
C2-C1-O3	119.4
C2-C1-H4	119.4
O3-C1-H4	121.2
C1-C2-S5	108.7
C1-C2-O6	103.1
S5-C2-O6	112.3
O3-Mg10-O6	81.3
C1-C2-H7	111.0
S5-C2-H7	114.1
O6-C2-H7	107.1
C2-S5-H8	96.0
C2-O6-H9	111.9
C1-O3-Mg10	119.4
C2-O6-Mg10	116.8
H9-O6-Mg10	131.3
O3-C1-C2-S5	119.3
O3-C1-C2-H7	-114.3

Table 6-13 Optimised Geometry Of Molecule B Plus Mg^{2+}, In Angstroms And Degrees.

	3-21G*
C1-C2	1.518
C1-O3	1.249
C1-H4	1.076
C2-S5	1.837
C2-O6	1.407
C2-H7	1.084
S5-H8	1.325
O3-Mg9	1.948
O6-Mg9	1.793
C2-C1-O3	118.8
C2-C1-H4	121.3
O3-C1-H4	119.9
C1-C2-S5	105.8
C1-C2-O6	109.1
S5-C2-O6	112.6
O3-Mg9-O6	87.5
C1-C2-H7	108.2
S5-C2-H7	108.7
O6-C2-H7	112.2
C2-S5-H8	94.2
C1-O3-Mg9	110.2
C2-O6-Mg9	114.4
O3-C1-C2-S5	121.3
O3-C1-C2-H7	-122.3

Table 6-14 Optimised Geometry Of Molecule C Plus Mg^{2+},

In Angstroms And Degrees.

	3-21G*
C1-C2	1.321
C1-O3	1.365
C1-H4	1.070
C2-S5	1.768
C2-O6	1.493
S5-H7	1.324
O6-H8	0.966
O3-Mg9	1.799
O6-Mg9	1.910
C2-C1-O3	121.4
C2-C1-H4	121.1
O3-C1-H4	117.4
C1-C2-S5	128.6
C1-C2-O6	112.8
S5-C2-O6	118.6
O3-Mg9-O6	88.8
C2-S5-H7	96.7
C2-O6-H8	114.4
C1-O3-Mg9	110.0
C2-O6-Mg9	106.0
H8-O6-Mg9	139.6

Table 6-15 Optimised Geometry Of Molecule D Plus Mg^{2+},

In Angstroms And Degrees.

	3-21G*
C1-C2	1.326
C1-O3	1.411
C1-H4	1.069
C2-S5	1.774
C2-O6	1.401
S5-H7	1.323
O3-Mg8	1.813
O6-Mg8	1.821
C2-C1-O3	120.0
C2-C1-H4	123.8
O3-C1-H4	116.2
C1-C2-S5	124.3
C1-C2-O6	120.7
O3-Mg8-O6	98.3
S5-C2-O6	115.0
C2-S5-H7	94.1
C1-O3-Mg8	100.6
C2-O6-Mg8	100.4

Table 6-16 Optimised Geometry Of Molecule E Plus Mg^{2+},

In Angstroms And Degrees.

	3-21G*
C1-C2	1.313
C1-O3	1.452
C1-H4	1.067
C2-S5	1.761
C2-O6	1.455
S5-H7	1.325
O6-H8	0.974
O3-H9	0.975
O3-Mg10	1.894
O6-Mg10	1.916
C2-C1-O3	117.1
C2-C1-H4	126.7
O3-C1-H4	116.2
C1-C2-S5	125.1
C1-C2-O6	114.0
S5-C2-O6	120.8
O3-Mg10-O6	84.7
C2-S5-H7	96.1
C2-O6-H8	114.7
C1-O3-H9	112.9
C1-O3-Mg10	111.8
H9-O3-Mg10	135.3
C2-O6-Mg10	112.5
H8-O6-Mg10	132.9

Table 6-17 Optimised Geometry Of Molecule F Plus Mg^{2+},

In Angstroms And Degrees.

	3-21G*
C1-C2	1.325
C1-O3	1.496
C1-H4	1.062
C2-S5	1.757
C2-O6	1.362
S5-H7	1.323
O3-H8	0.966
O3-Mg9	1.890
O6-Mg9	1.808
C2-C1-O3	113.9
C2-C1-H4	131.3
O3-C1-H4	114.9
C1-C2-S5	122.5
C1-C2-O6	120.2
S5-C2-O6	117.3
O3-Mg9-O6	90.3
C2-S5-H7	93.6
C1-O3-H8	112.5
C1-O3-Mg9	105.5
H8-O3-Mg9	142.0
C2-O6-Mg9	110.1

Table 6-18 Optimised Geometry Of Molecule G Plus Mg^{2+},

In Angstroms And Degrees.

	3-21G*
C1-C2	1.544
C1-O3	1.414
C1-H4	1.085
C2-S5	1.705
C2-O6	1.268
C1-H7	1.085
S5-H8	1.327
O3-Mg9	1.790
O6-Mg9	1.918
C2-C1-O3	108.8
C2-C1-H4	107.9
O3-C1-H4	112.1
C1-C2-S5	119.9
C1-C2-O6	117.7
S5-C2-O6	122.3
O3-Mg9-O6	88.9
C2-C1-H7	107.9
O3-C1-H7	112.1
H4-C1-H7	107.9
C2-S5-H8	95.8
C1-O3-Mg9	114.0
C2-O6-Mg9	110.6
O6-C2-C1-H4	121.8
O6-C2-C1-H7	-121.8

Table 6-19 Optimised Geometry Of Molecule H Plus Mg^{2+},

In Angstroms And Degrees.

	3-21G*
C1-C2	1.528
C1-O3	1.492
C1-H4	1.081
C2-S5	1.682
C2-O6	1.268
C1-H7	1.081
S5-H8	1.329
O3-H9	0.975
O3-Mg10	1.929
O6-Mg10	1.870
C2-C1-O3	105.4
C2-C1-H4	111.3
O3-C1-H4	108.7
C1-C2-S5	118.6
C1-C2-O6	116.6
S5-C2-O6	124.9
O3-Mg10-O6	82.7
C2-C1-H7	111.3
O3-C1-H7	108.7
H4-C1-H7	111.1
C2-S5-H8	95.4
C1-O3-H9	111.8
C1-O3-Mg10	115.2
C2-O6-Mg10	120.0
H9-O3-Mg10	133.0
O6-C2-C1-H4	117.7
O6-C2-C1-H7	-117.7

Table 6-20 Optimised Geometry Of Molecule A, Ethyl Replacing Thiol Hydrogen, In Angstroms And Degrees.

	3-21G*
C1-C2	1.515
C1-O3	1.209
C1-H4	1.081
C2-S5	1.821
C2-O6	1.421
C2-H7	1.082
S5-C8	1.822
O3-H9	2.071
O6-H9	0.972
C8-H10	1.080
C8-C12	1.543
C12-H13	1.083
C2-C1-O3	121.0
C2-C1-H4	116.2
O3-C1-H4	122.9
C1-C2-S5	107.9
C1-C2-O6	109.8
S5-C2-O6	111.2
C1-C2-H7	109.4
S5-C2-H7	108.7
O6-C2-H7	109.8
C2-S5-C8	98.0
C2-O6-H9	109.0
S5-C8-H10	108.3
H10-C8-H11	110.6
S5-C8-C12	109.2
H10-C8-C12	110.2
C8-C12-H13	110.4
H13-C12-H14	108.5
O3-C1-C2-S5	120.5
O3-C1-C2-H7	-124.4

Table 6-21 Optimised Geometry Of Molecule C, Methyl Replacing Thiol Hydrogen, In Angstroms And Degrees.

	3-21G*
C1-C2	1.338
C1-O3	1.304
C1-H4	1.094
C2-S5	1.782
C2-O6	1.434
S5-C7	1.821
O3-H8	2.013
O6-H8	0.980
C7-H9	1.081
C7-H10	1.081
C7-H11	1.081
C2-C1-O3	122.7
C2-C1-H4	116.6
O3-C1-H4	120.7
C1-C2-S5	128.1
C1-C2-O6	115.6
S5-C2-O6	116.3
C2-S5-C7	100.9
C2-O6-H8	99.8
S5-C7-H9	109.1
S5-C7-H10	109.1
H9-C7-H10	109.8

Table 6-22 Optimised Geometry Of Molecule E, Methyl Replacing Thiol Hydrogen, In Angstroms And Degrees.

	3-21G*
C1-C2	1.313
C1-O3	1.410
C1-H4	1.067
C2-S5	1.769
C2-O6	1.379
S5-C7	1.822
O3-H8	2.109
O6-H8	0.970
O3-H9	0.961
C7-H10	1.079
C7-H11	1.079
C7-H12	1.079
C2-C1-O3	116.0
C2-C1-H4	124.2
O3-C1-H4	119.8
C1-C2-S5	123.2
C1-C2-O6	120.6
S5-C2-O6	116.3
C2-S5-C7	102.1
C2-O6-H8	109.0
C1-O3-H9	112.7
S5-C7-H10	108.9
H10-C7-H11	110.0

Table 6-23 Optimised Geometry Of Molecule F, Methyl Replacing Thiol Hydrogen, In Angstroms And Degrees.

	3-21G*
C1-C2	1.332
C1-O3	1.439
C1-H4	1.065
C2-S5	1.892
C2-O6	1.278
S5-C7	1.810
O3-H8	0.970
O6-H8	2.278
C7-H9	1.083
C7-H10	1.083
C7-H11	1.083
C2-C1-O3	118.3
C2-C1-H4	126.9
O3-C1-H4	114.8
C1-C2-S5	114.1
C1-C2-O6	128.3
S5-C2-O6	117.6
C2-S5-C7	97.3
C1-O3-H8	103.7
S5-C7-H9	108.6
H9-C7-H10	110.3

Table 6-24 Optimised Geometry Of Molecule H, Ethyl Replacing Thiol Hydrogen, In Angstroms And Degrees.

	3-21G*
C1-C2	1.523
C1-O3	1.425
C1-H4	1.083
C2-S5	1.752
C2-O6	1.214
C1-H7	1.083
S5-C8	1.824
O3-H9	0.971
O6-H9	2.058
C8-H10	1.078
C8-H11	1.078
C8-C12	1.543
C12-H13	1.083
C12-H14	1.083
C12-H15	1.083
C2-C1-O3	109.6
C2-C1-H4	110.5
O3-C1-H4	110.5
C1-C2-S5	116.0
C1-C2-O6	120.5
S5-C2-O6	123.5
C2-C1-H7	108.9
O3-C1-H7	110.5
H4-C1-H7	108.4
C2-S5-C8	99.3
C1-O3-H9	108.6
S5-C8-H10	108.6
S5-C8-C12	108.7
H10-C8-C12	110.3
C8-C12-H13	110.5
H13-C12-H14	108.4
H4-C1-C2-O6	121.0
H7-C1-C2-O6	-121.0

Figure 6-1 The Model Reaction Scheme.

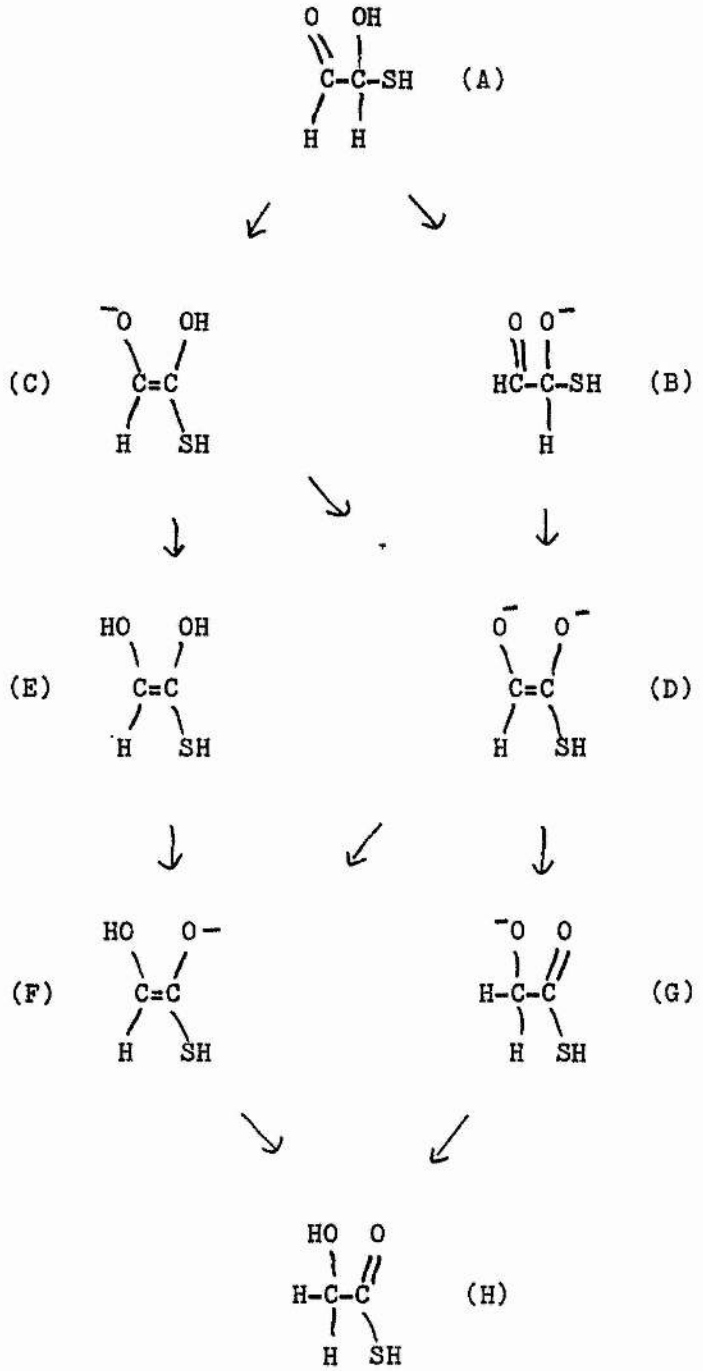
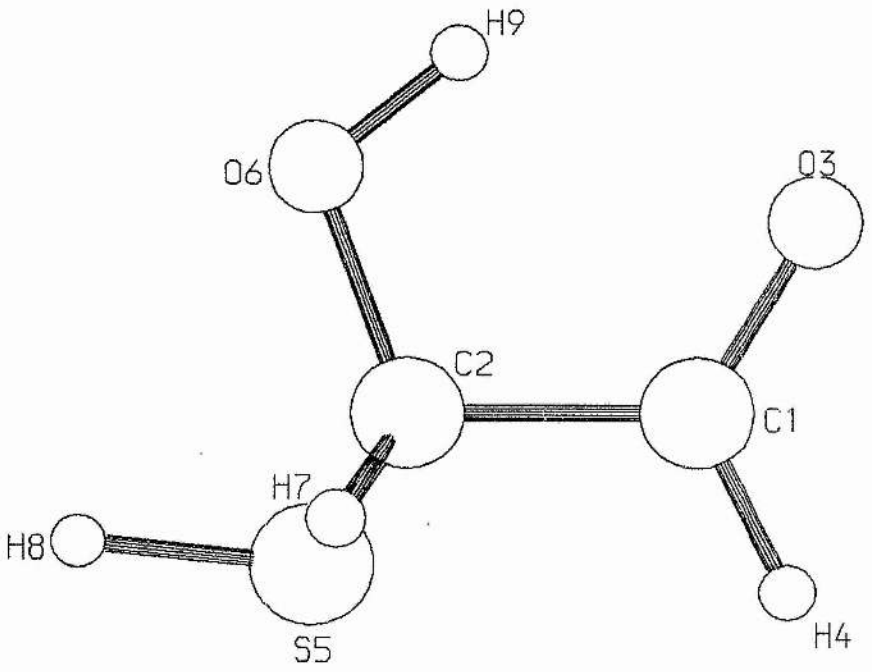


Figure 6-2 The 3-21G* Optimised Geometry Of Molecule A.



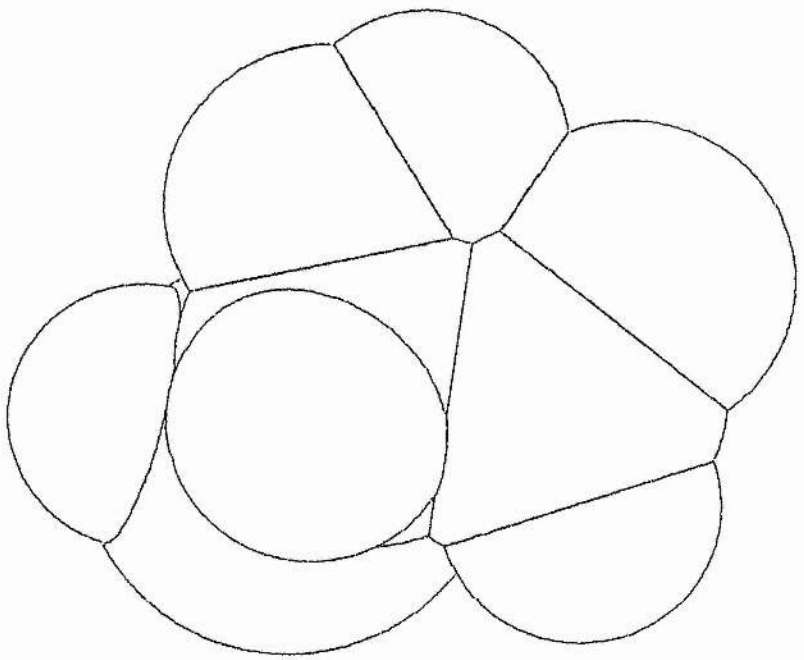
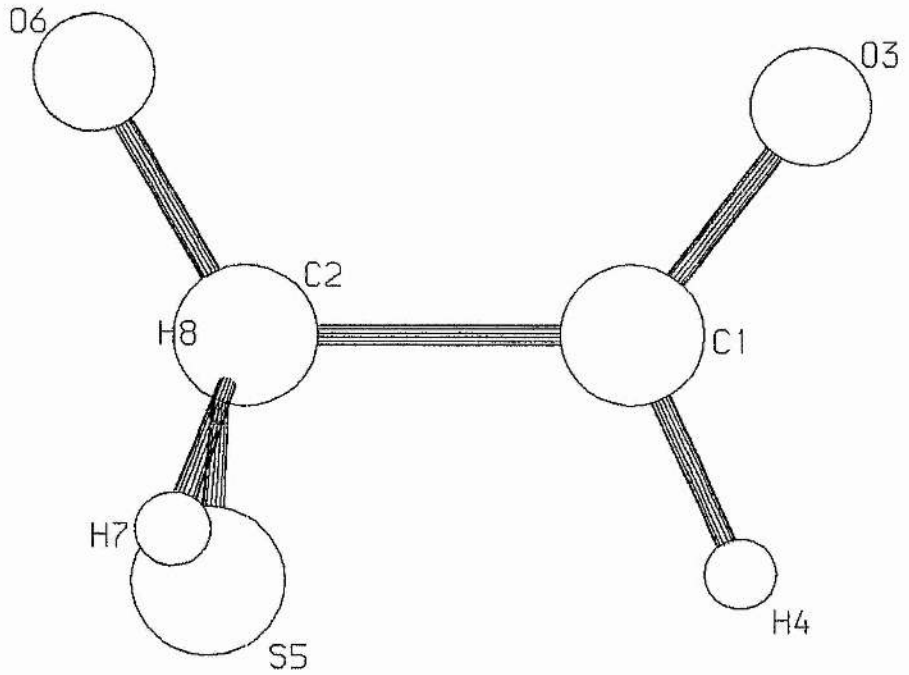


Figure 6-3 The STO-3G Optimised Geometry Of Molecule B.



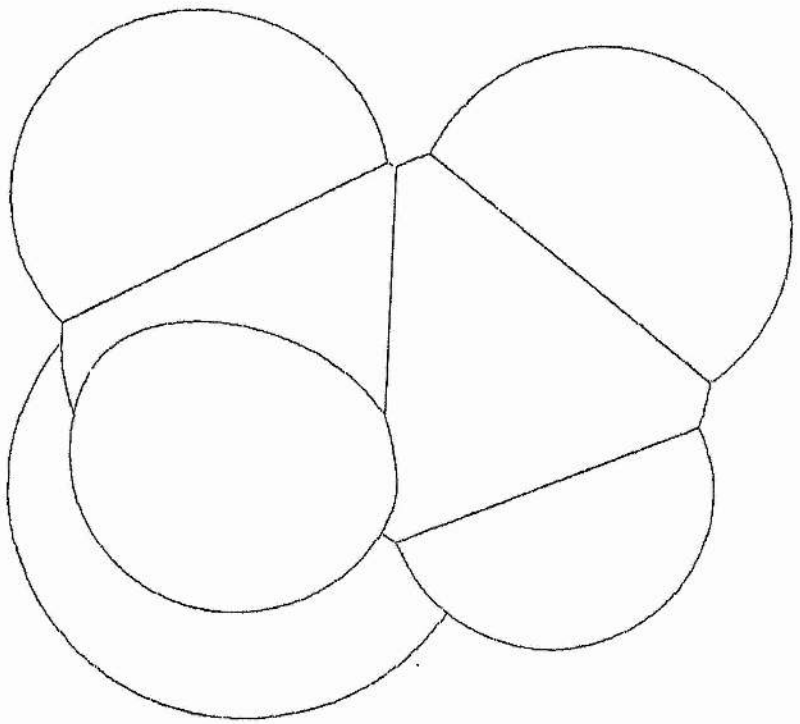
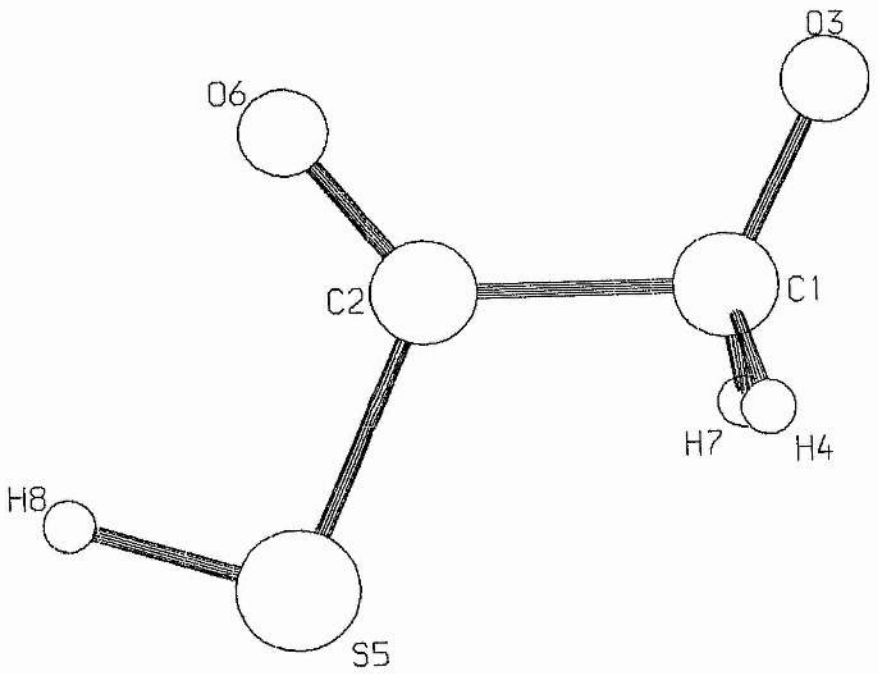


Figure 6-4 The STO-3G Optimised Geometry Of Molecule G.



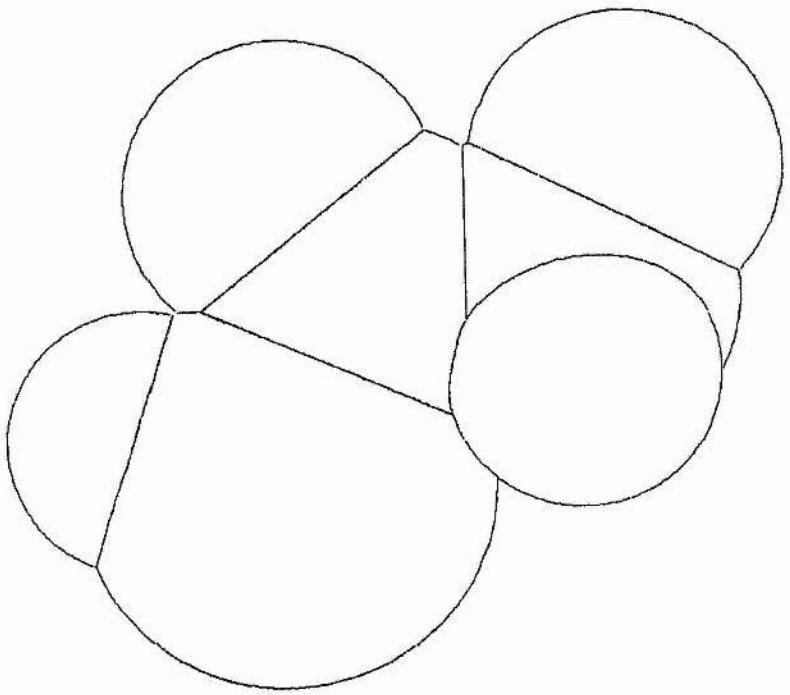
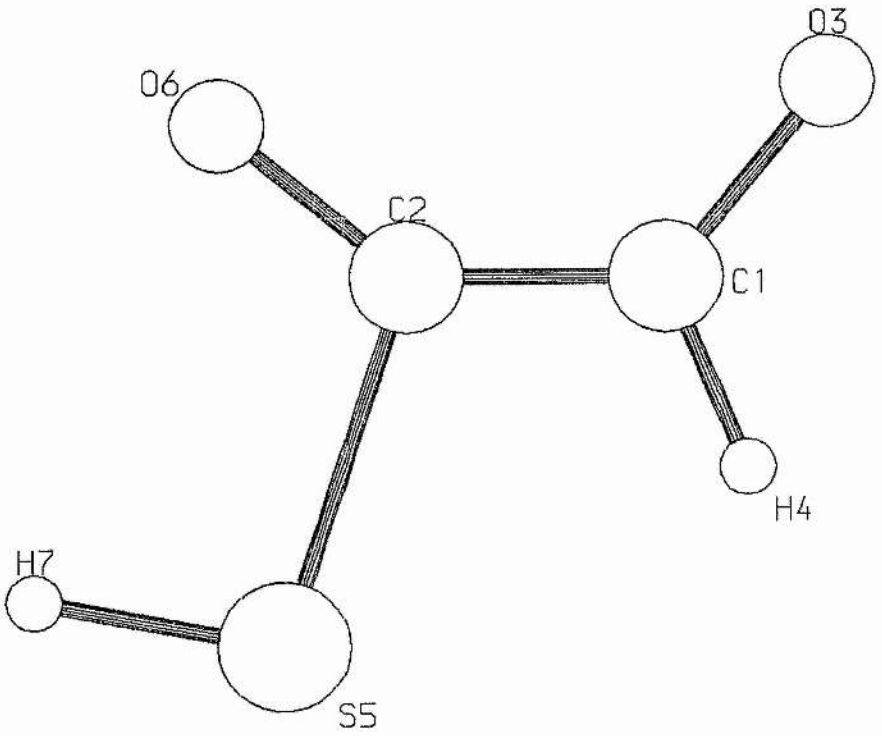


Figure 6-5 The STO-3G Optimised Geometry Of Molecule D.



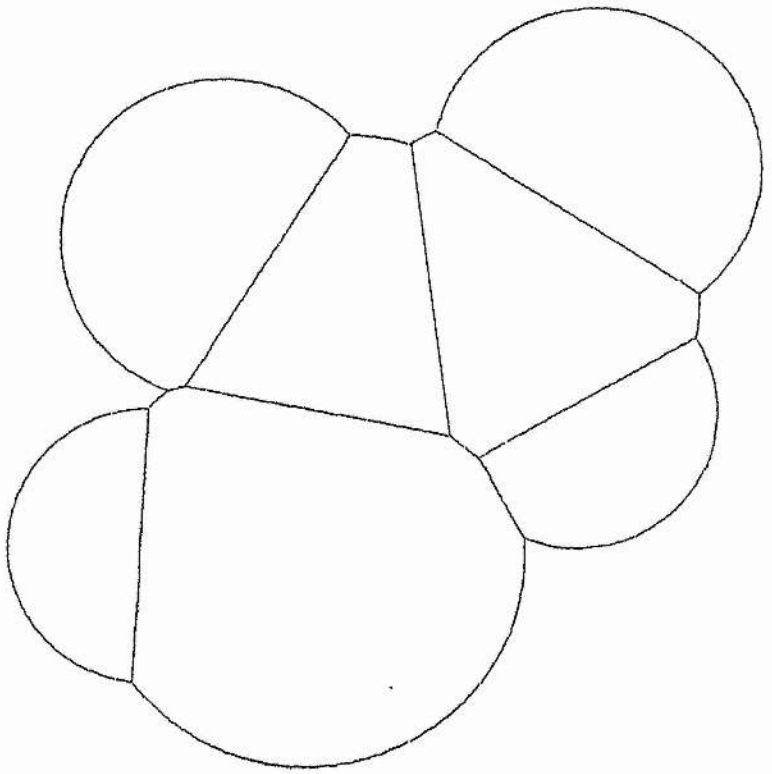
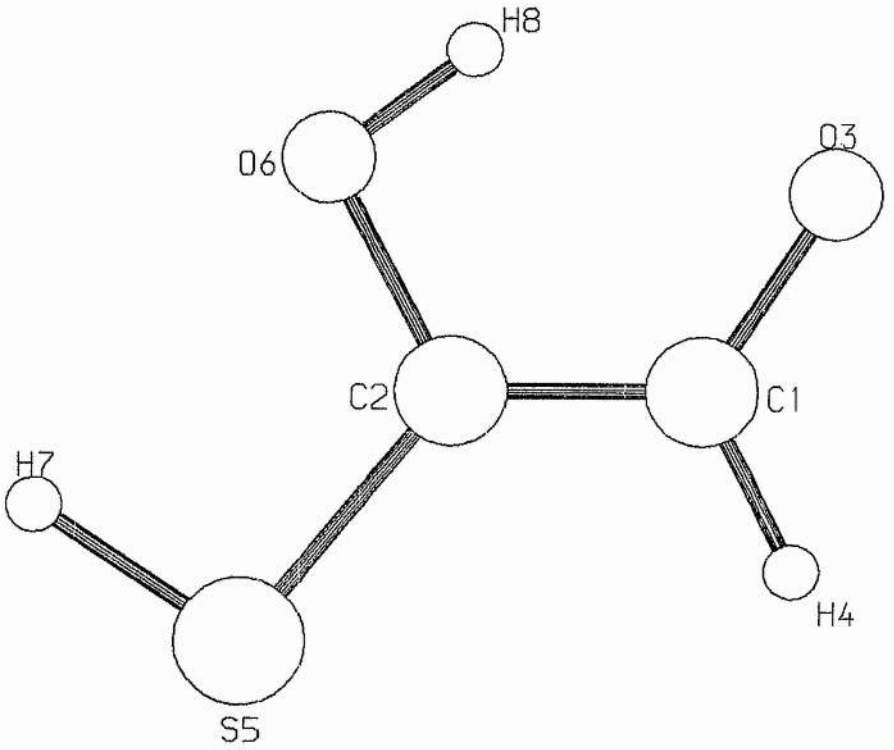


Figure 6-6 The 3-21G* Optimised Geometry Of Molecule C.



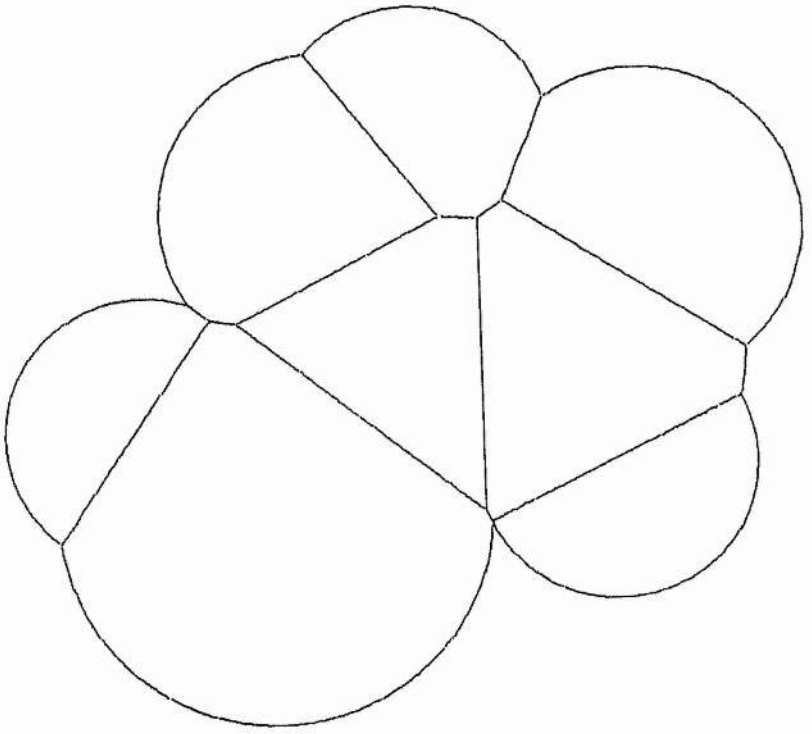
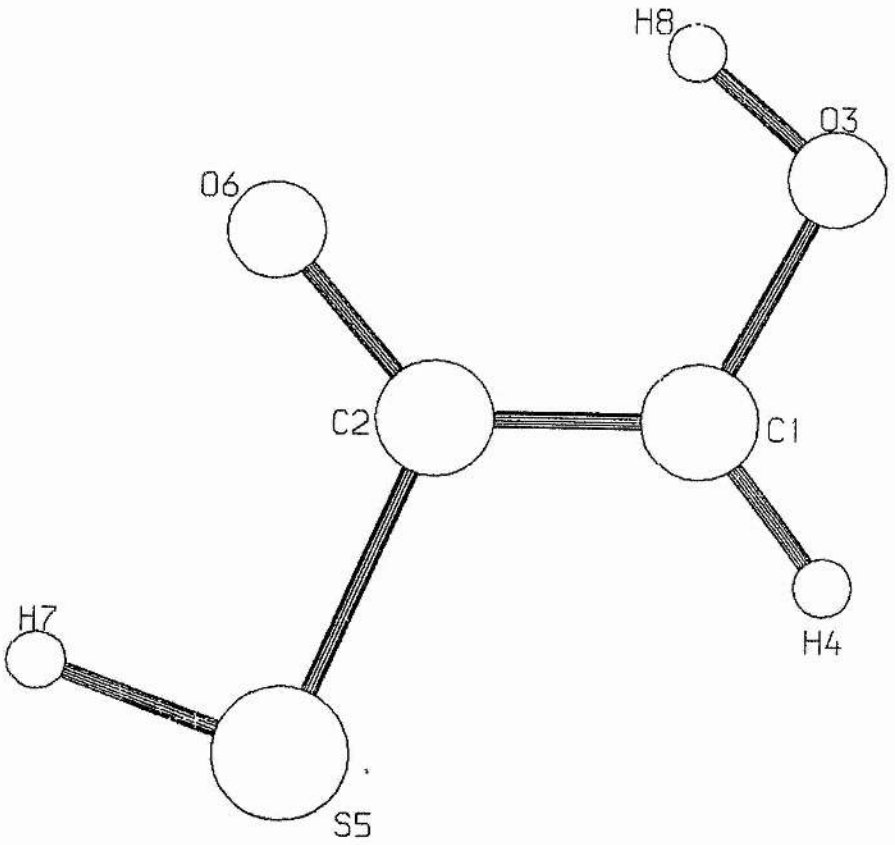


Figure 6-7 The 3-21G* Optimised Geometry Of Molecule F.



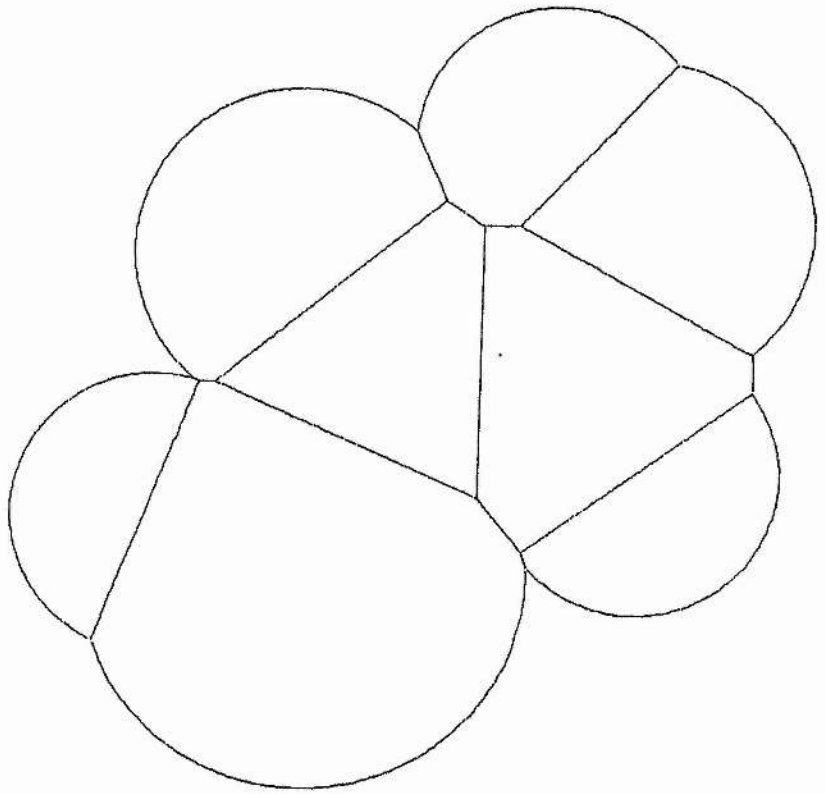
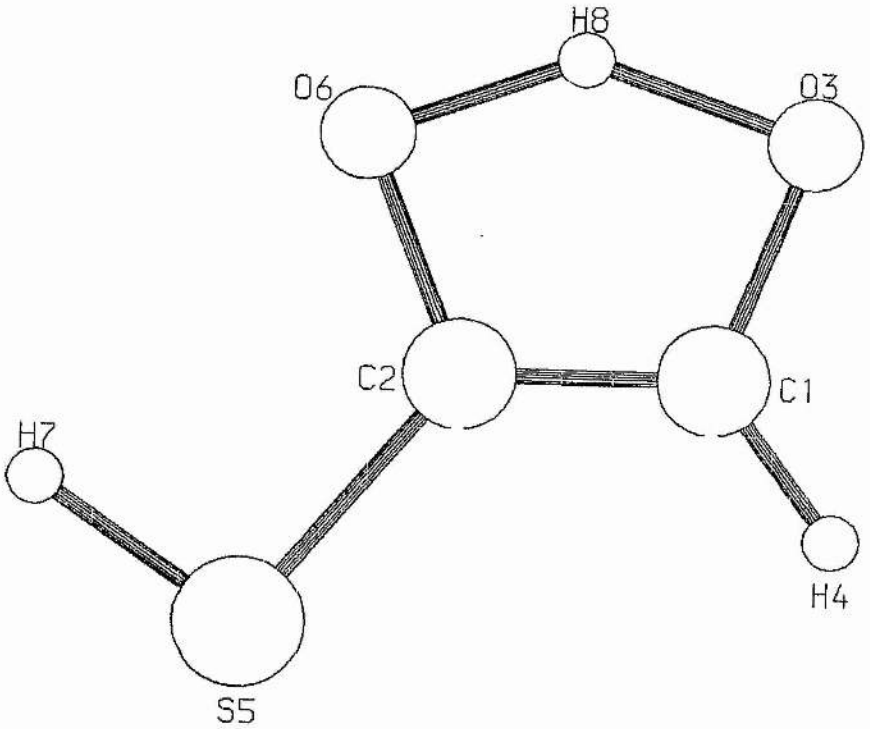


Figure 6-8 The 3-21G* Optimised Geometry Of Molecule CF.



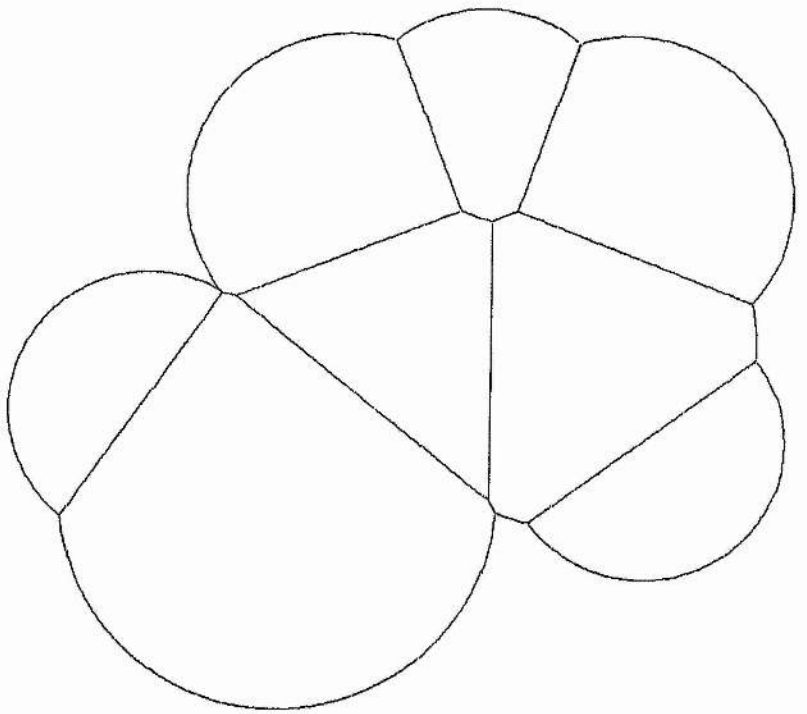
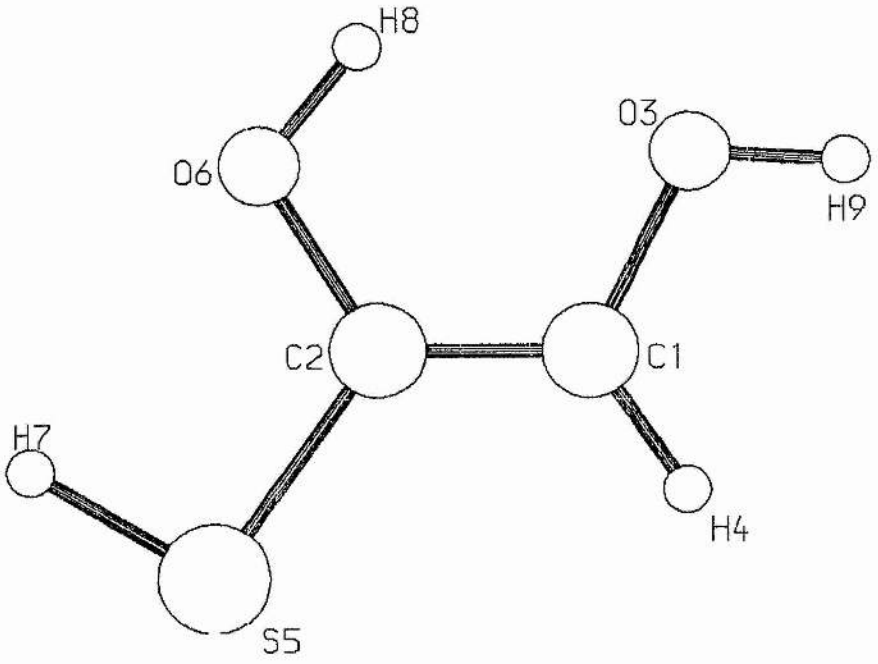


Figure 6-9 The 3-21G* Optimised Geometry Of Molecule E.



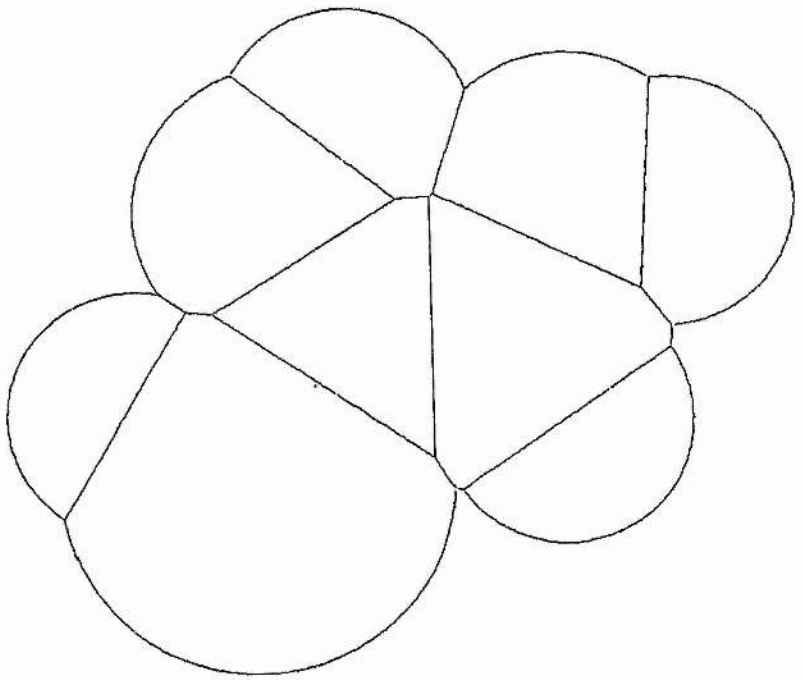
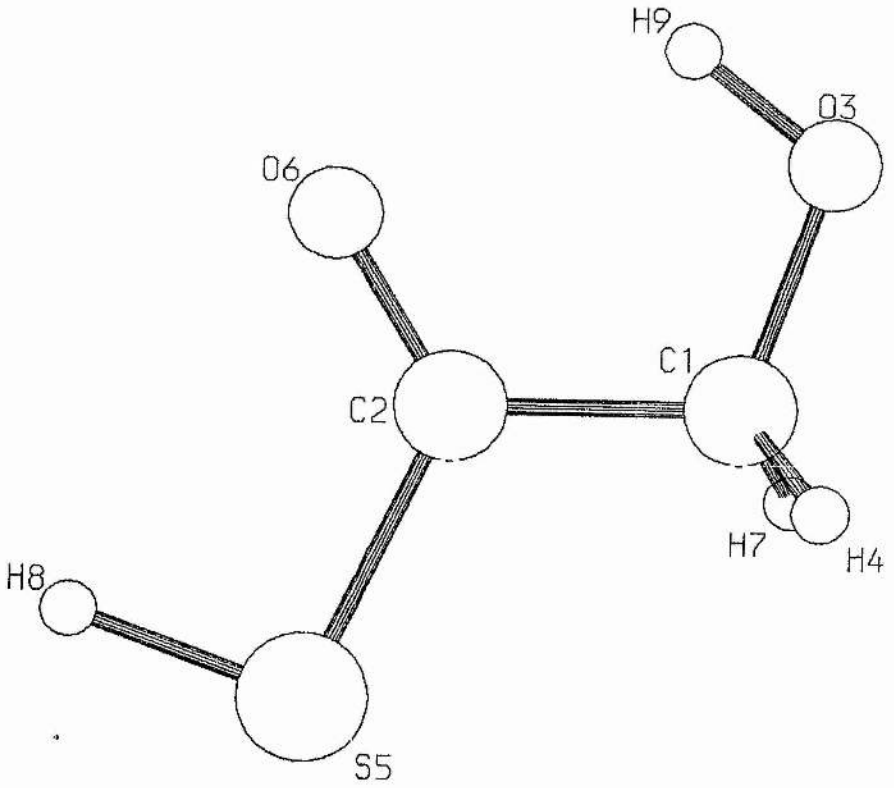


Figure 6-10 The 3-21G* Optimised Geometry Of Molecule H.



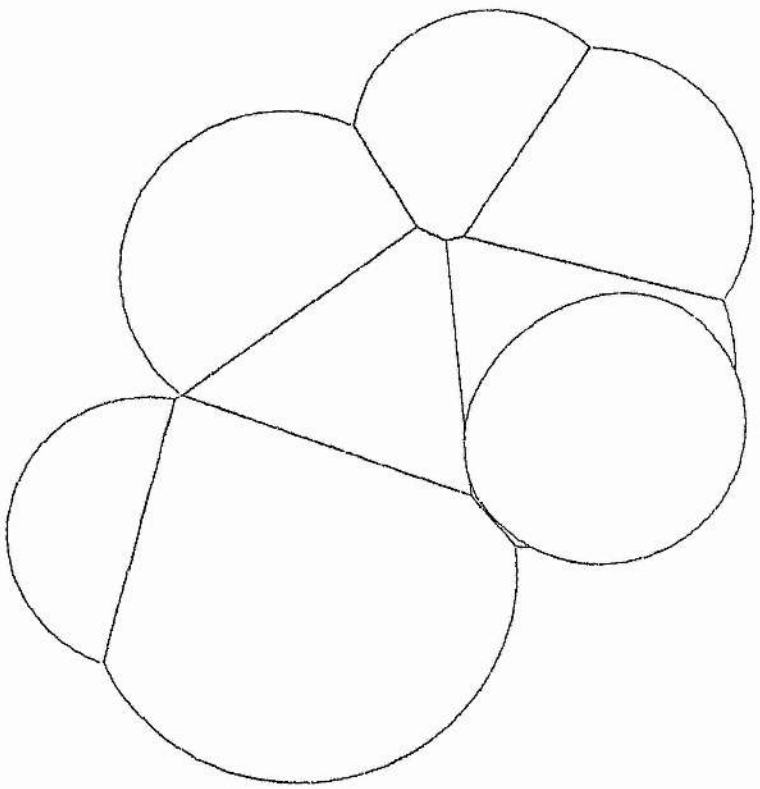


Figure 6-11 The 3-21G* Optimised Geometry Of Molecule A Plus Mg^{2+}.

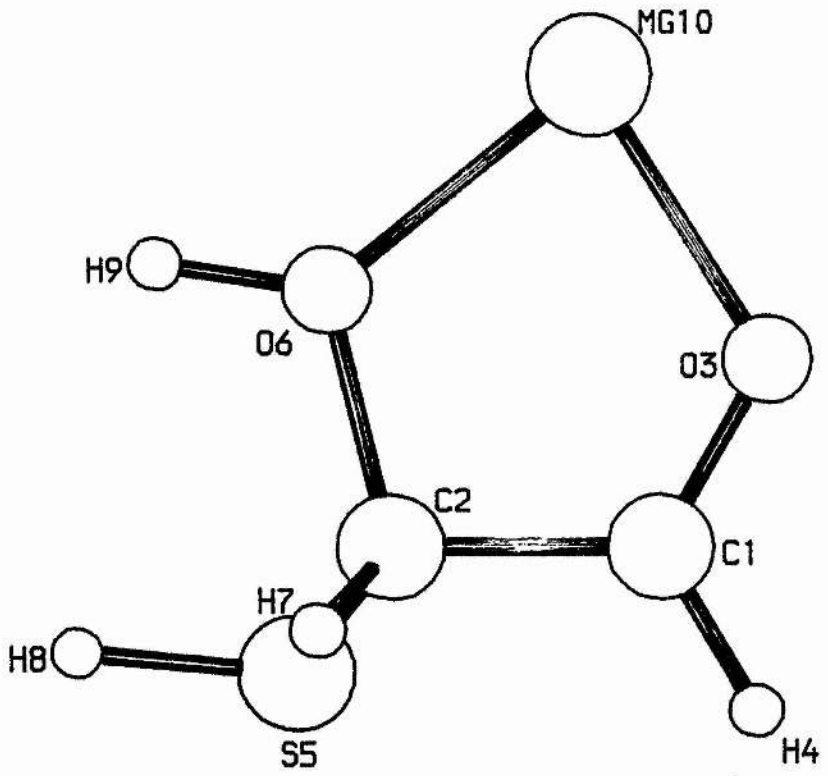


Figure 6-12 The 3-21G* Optimised Geometry Of Molecule B Plus Mg^{2+}.

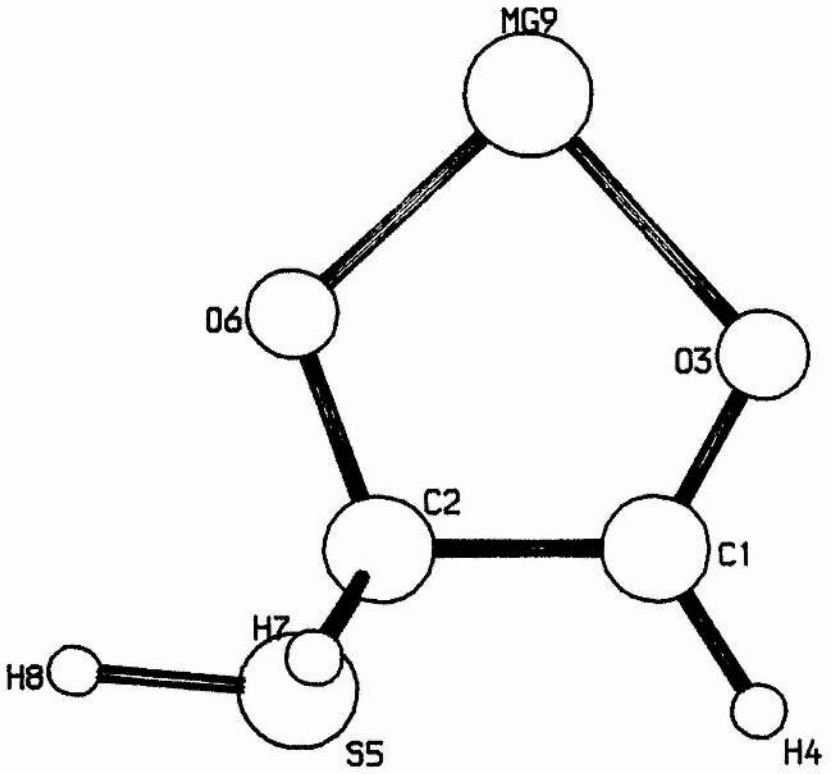


Figure 6-13 The 3-21G* Optimised Geometry Of Molecule C Plus Mg^{2+}.

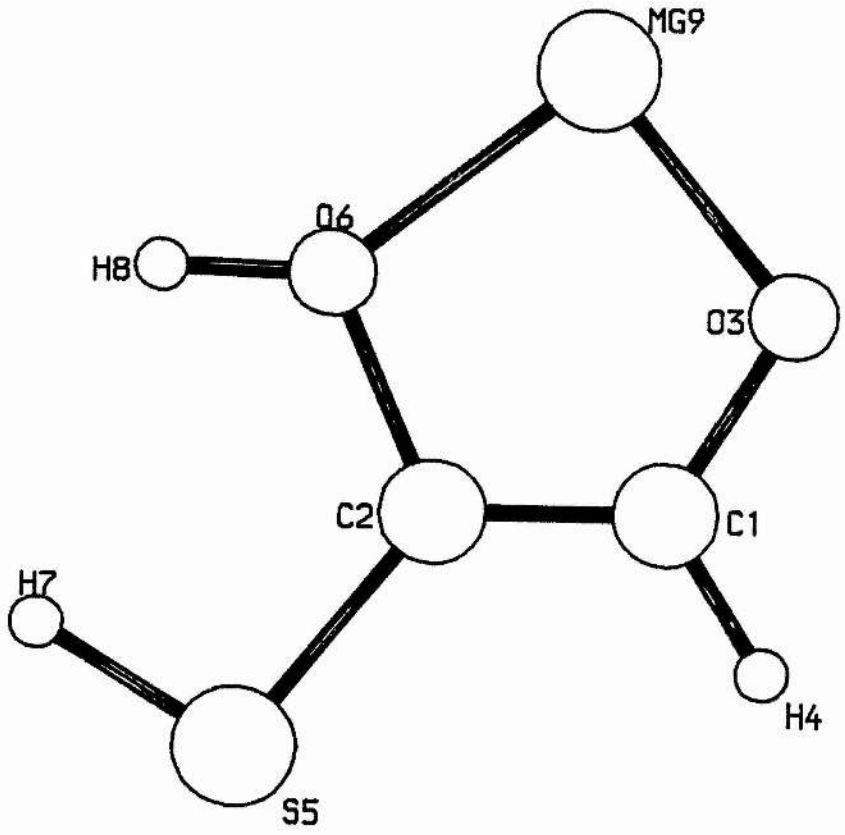


Figure 6-14 The 3-21G* Optimised Geometry Of Molecule D Plus Mg^{2+}.

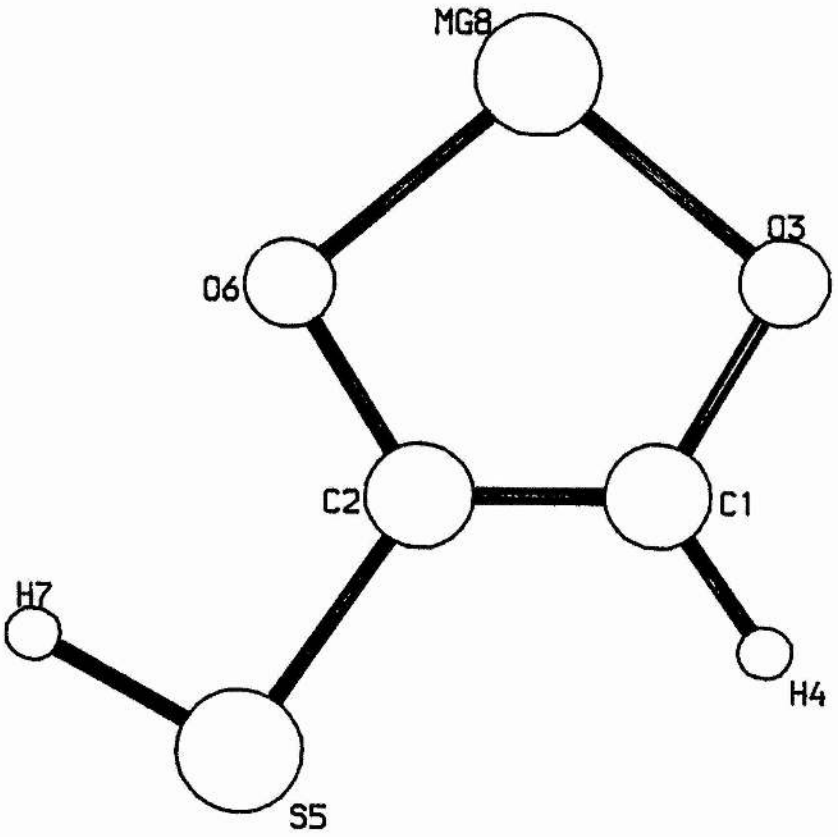


Figure 6-15 The 3-21G* Optimised Geometry Of Molecule E Plus Mg^{2+}.

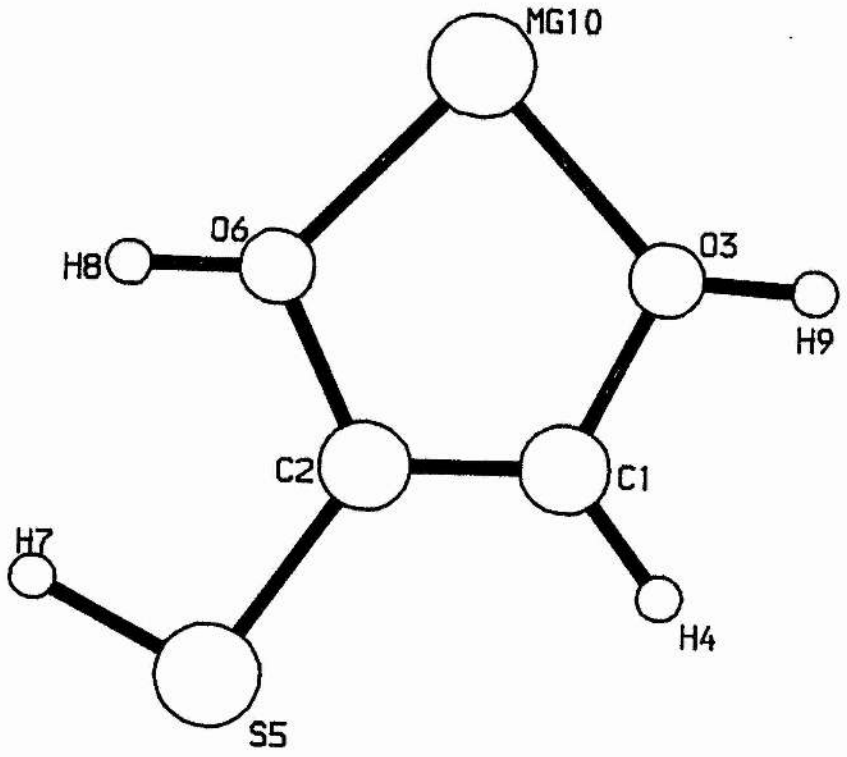


Figure 6-16 The 3-21G* Optimised Geometry Of Molecule F Plus Mg^{2+}.

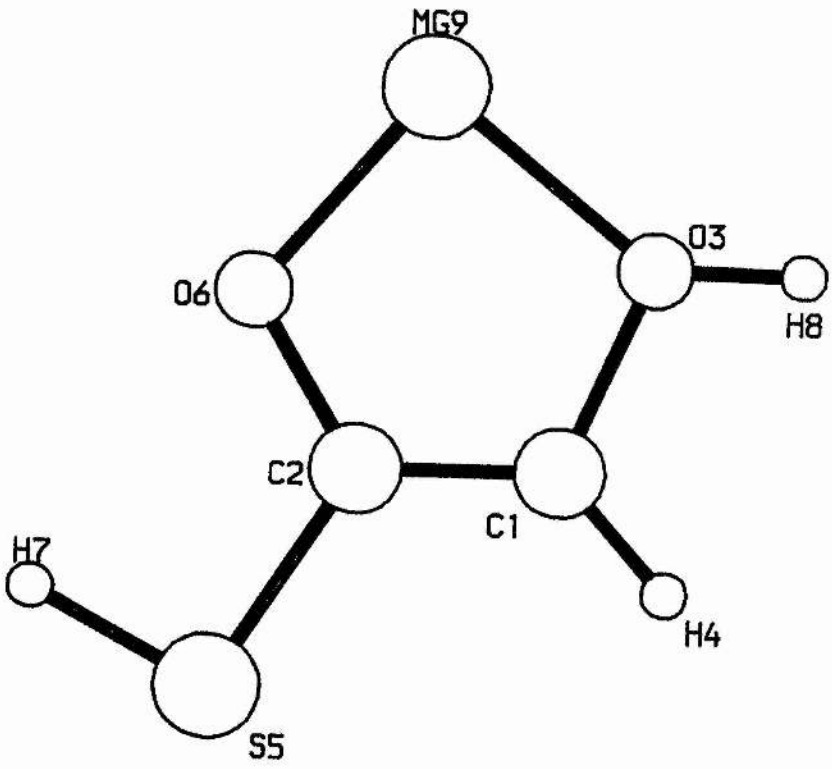


Figure 6-17 The 3-21G* Optimised Geometry Of Molecule G Plus Mg^{2+}.

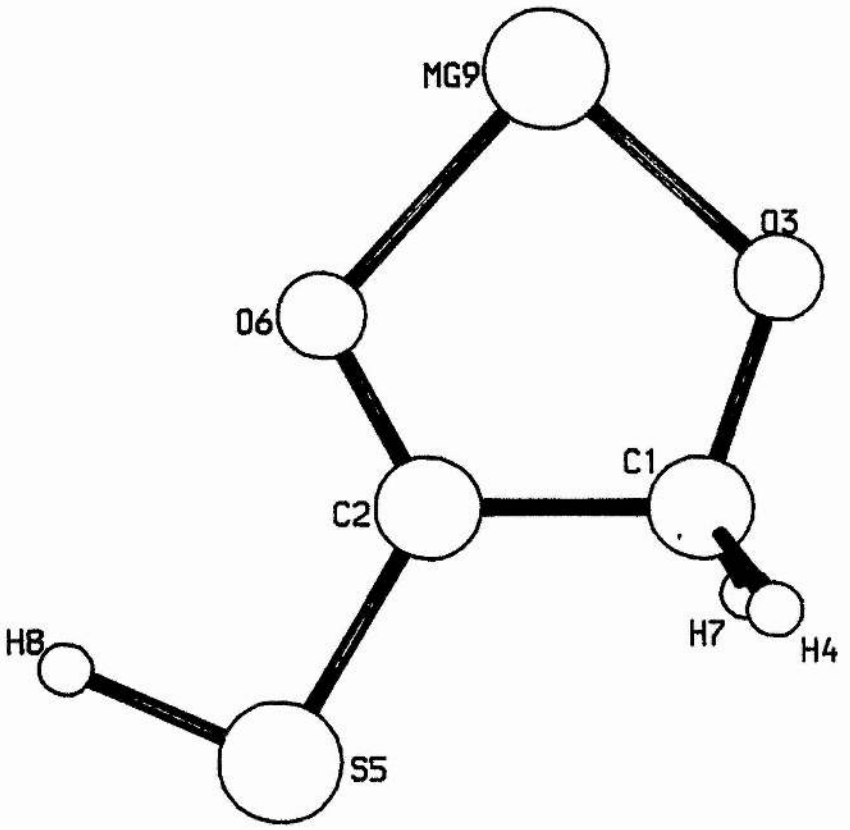
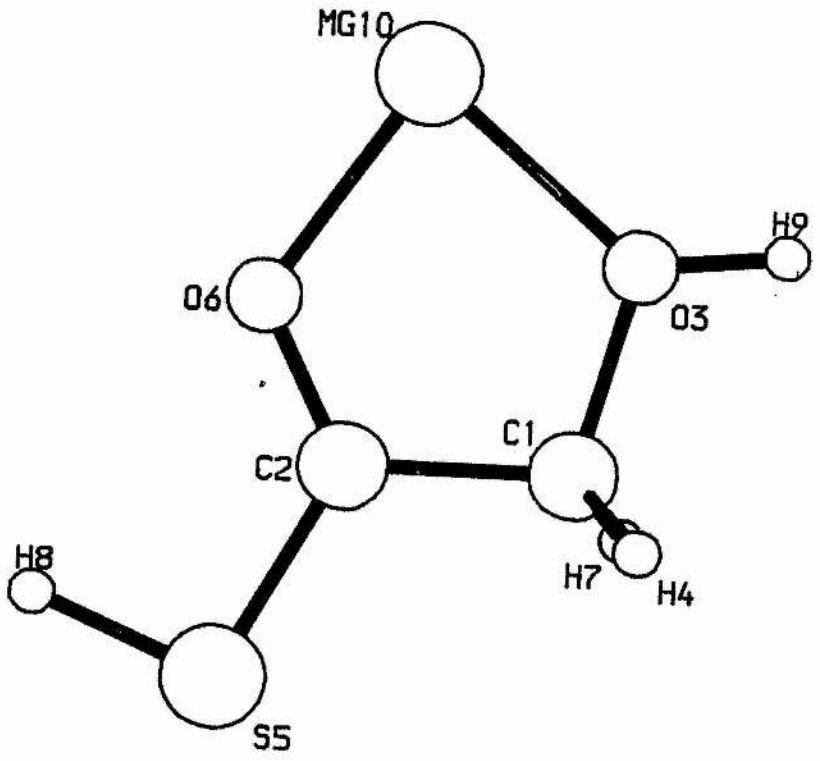


Figure 6-18 The 3-21G* Optimised Geometry Of Molecule H Plus Mg^{2+}.



CHAPTER 7

THE INHIBITORS OF GLYOXALASE I

The inhibition of glyoxalase I can be divided into the broad categories described in Chapter 3: namely glutathione-based inhibitors, natural inhibitors, "reductones" and mechanism-based inhibitors. This work concerns itself with the last category for the following reasons:

The design of more powerful inhibitory molecules for glyoxalase I could be achieved by an examination of the properties of the current mechanism-based inhibitors, providing there is some knowledge of the enzyme mechanism. This investigation could also work in reverse, as the inhibitors themselves provide information about the enzyme binding site properties. The natural inhibitors, on the other hand, do not lend themselves to the systematic development of other inhibitory species; they are usually present as a single chemical species in a biological system, and their structure can be as much due to the requirements and idiosyncracies of the organism's biosynthetic mechanisms as to the functional inhibition of glyoxalase I. The natural inhibitors can be good starting materials for the design of enzyme inhibitors, but, in the case of glyoxalase I, there is little point in examining these esoteric species in any great detail. This is because better inhibitors which are almost certainly easier to relate to the enzyme mechanism are available. The "reductones" have the paene-enediol structure, emphasised by Douglas and Nadvi [102], and so could be fitted into the mechanism-based inhibitor category, except that they were originally tried as inhibitors because of their anti-tumour properties [101].

The design of enzyme inhibitors - or indeed any drug - by rational approaches has a much greater appeal to any serious scientist than the screening of vast numbers of compounds in the hope that one might prove to be of use. In order to design enzyme inhibitors, it is necessary to have some knowledge of the enzyme reaction and, preferably, of the actual enzyme-substrate interactions. The relevant information available to us on the glyoxalase I reaction is presented in Chapters 3 and 6: to summarise, the enzyme seems certain to catalyse a proton transfer reaction as opposed to a hydride transfer, and this reaction is intermolecular, rather than intramolecular. A representation of the reaction is given in Figure 3-1. the active site of the enzyme must accommodate not just the glyoxal section of the substrate, which is the part which is chemically altered, but also the glutathione part. This is obvious when one considers that glutathione analogues are inhibitors of the enzyme-catalysed reaction. There is also evidence that there is a large hydrophobic region near the enzyme site [89,90,97].

It seems likely that the glutathione structure is responsible for the initial binding, placing the α-keto end of the substrate in the catalytic region of the enzyme. Therefore an inhibitor could either block the initial binding, as the glutathione analogues do, or the subsequent interaction with the catalytic apparatus, as the "transition-state" inhibitors are expected to do.

The mechanism-based inhibitors first described were those of Douglas and Nadvi [102], who found that maltol, 3-hydroxy-2-methyl-4H-pyran-4-one, and other paene-enediol structured compounds were inhibitors of glyoxalase I. This was followed by the publication of details of other "transition-state" inhibitors which were purported to resemble the enediol structure which is probably

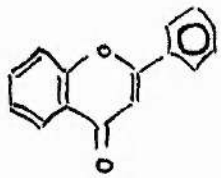
present at some stage in the enzyme-catalysed reaction. A model of the enediol intermediate is discussed in Chapter 6. It is a stable, neutral species of comparable energy to the substrate and product molecules; it is not a transition state species. The inhibitors are essentially intermediate-analogues, being related to the enediol, as opposed to the higher energy anionic species also investigated in Chapter 6.

The inhibitor molecules studied in this work are listed in Table 7-1, along with their I_{50} concentrations for the inhibition of glyoxalase I. These I_{50} values were obtained from Professor Richard Brandt and his co-workers, using a system including human erythrocyte glyoxalase I, which is probably the nearest currently used *in vitro* system to *in vivo* glyoxalase I. The corresponding structures for these compounds are shown in Figures 7-1 and 7-2.

The original theoretical studies of this group on the inhibitors were carried out prior to the discovery that the flavones and most of the coumarins derivatives were active. Some of these less inhibitory compounds have been studied in this work, but the majority of effort has been directed at investigating the flavones, and subsequently the coumarin derivatives. This is for two main reasons: the flavones and coumarins are better inhibitors than the other molecules and they form two groups of structurally-related compounds which lend themselves to systematic study.

7.1 The Flavonoid Compounds

The flavonoid compounds compounds studied were mainly flavonols, which are oxidised forms of a group of chemicals, the flavones, which possess the structure below [163]:



The A and B rings behave as aromatic rings, particularly the hydroxyflavones which resemble phenols in most of their reactions, and the C ring behaves as a modified pyrone ring. Flavonols are hydroxy derivatives of the basic flavone shown above. Many of them are found in nature [163], occurring in higher plants, but not in animals. They usually occur as glycosides, which can be easily hydrolysed back to the respective flavonols. The characteristic oxidation pattern found for most flavonols, with hydroxy or methoxy groups at positions 3, 5 and 7 and at any of 3', 4' and 5', is due to the mixed biosynthesis of these compounds: ring A is derived from acetate, but ring B and the carbons of ring C are the products of the shikimic acid pathway [164]. This results in the formation of four basic patterns of flavonol structure, based on the extent of hydroxylation of ring B: the simplest is galangin, 3,5,7-trihydroxyflavone, then in order of extent of hydroxylation, kaempferol, 3,5,7,4'-tetrahydroxyflavone, quercetin, 3,5,7,3',4'-pentahydroxyflavone, and myricetin, 3,5,7,3',4',5'-hexahydroxyflavone. Various other natural flavonols are created from these basic ones by methylation, or addition of extra hydroxyl groups.

All the flavonols and their glycosides are yellow or pale yellow, but they are not responsible for yellow coloration in flowers as they are present in too small a concentration. The yellow coloration is generally due to carotenoid pigments. The flavonols do give body to white flowers, which would otherwise appear transparent, and their ultraviolet absorption is probably important in attracting insects to the plants. Flavonols are also found associated with the lignin of wood and leaves [163].

The flavonoid family are responsible for a wide range of biological effects, which include the inhibition of various enzymes. These include aldose reductase [165,166], lipoxygenase [167,168], aromatase [169], Ca^{2+}/phospholipid-dependent protein kinase [170], phosphorylase kinase [171] and tyrosine protein kinase [171]. Some flavones have been shown to inhibit glyoxalase I very effectively [107].

The flavonoids have also been suggested as antiviral agents, based on the activity of certain 3-methoxyflavones against human rhinovirus, which is considered to be the main cause of the common cold [172]. They have also been proposed for the treatment of cataracts in diabetics [165,166], as they inhibit aldose reductase, an enzyme which seems to initiate the process leading to the formation of sugar cataracts.

Flavonols such as quercetin have also been found to be active in inhibiting the tumorigenesis of mouse skin tumours. This has been ascribed to various mechanisms.

Flavones, along with ellagic acid, another phenolic compound found in plants, were found to inhibit tumour formation in mouse skin which was treated initially with bay-region diol-epoxides and then with TPA [173]. This inhibition could be by direct action on the diol-epoxides [173,174] as these species disappear from solution when flavones are added to them *in vitro* [173], or the inhibition could be due to inhibition of enzymes important in tumorigenesis. This latter explanation was invoked by Kato et al. [167] to explain the observed inhibition of tumour promotion in mouse skin by TPA. They suggested that the flavones inhibited lipoxygenase. It was subsequently shown by Wheeler and Berry [168], that certain flavonoids do indeed inhibit lipoxygenase. An interesting effect of flavone, itself, on tumour development was noted by Alworth and Slaga [175]: the effect was found to be concentration dependent; high doses inhibited tumorigenesis, whereas low doses seemed to enhance it. This was suggested to be due to different rates of reaction on more than one enzyme or receptor. Another flavone, quercetin, was found to be a potent inhibitor of calcium and phospholipid-dependent protein kinase activity [170]. The flavonol also inhibits the enzyme when it is stimulated by TPA instead of calcium. The inhibition probably occurs at a different site from the TPA binding site, as quercetin does not appear to affect TPA binding.

The flavonoid compounds studied were mostly hydroxylated derivatives of 3-hydroxyflavone, with two derivatives of 3-hydroxyflavanone, naringenin and taxifolin. These flavonoids are shown in Figure 7-1. The flavone compounds are much more potent inhibitors than the flavanone ones, indicating that the double bond between carbon 2 and carbon 3 is important, or that the conformation of the molecule influenced by this bonding is important: that is, the

substituent groups on C2 and C3 are in the plane of the rings A and C in the flavones, but are at an angle to this plane when C2 and C3 are sp^3 centres. The inhibition increases with the extent of hydroxylation of the phenyl ring B, but this is not the sole determinant of inhibitory power, as 3-hydroxyflavone, with no hydroxyl groups on the phenyl ring, is as good an inhibitor as quercetin. On the other hand, the hydroxylation pattern at carbons 5 and 7 does not seem to be important, as 3-hydroxyflavone does not possess either of these groups and fisetin only has the 7-hydroxy group, but is of comparable strength to quercetin. Apart from these two, the other molecules studied do have the 5- and 7-hydroxy groups. This is a legacy from the acetate biosynthetic pathway, which is used to construct ring A.

It was not possible to carry out a full geometry optimisation of all the flavone compounds because of the constraints of available time imposed on the work. Nor is it completely necessary, as the property of interest, the molecular electrostatic potential, has proved to be insensitive to the quality of the wave function [135]. The basic skeleton of the molecules - the chromone double ring - was optimised using the STO-3G basis set, however. The results are listed in Table 7-2 and the optimised geometry is depicted in Figure 7-3.

The lowest energy form has the hydroxy group pointing towards the carbonyl group. The carbonyl group itself is distorted slightly towards the hydroxy side of the molecule as C2-C7-O11 is 126.1°, whereas the other angle, C8-C7-O11, is only 120.4°. The carbon-carbon bonding distance are all about 1.4 A, except for C2-C7 and C7-C8, which are about 1.5 A, and C4-C8 which is 1.325 A. These latter-mentioned deviations occur because the pyrone ring is not a delocalised structure, but has a definite double bond, C4-C8, and two single bonds, C2-C7 and C7-C8. This is manifested in the chemistry as well as the

bonding of this part of the compound.

Studies of the γ-pyrone ring itself, as well as the sulphur-replaced derivatives, have been carried out and published [176]. The γ-pyrone species was optimised using the 3-21G basis set. It was found that the bond lengths corresponding to C4-C8 and C1-C2 were 1.321 A in this basis, whereas those corresponding to C2-C7 and C7-C8 were 1.469 A. There was little evidence of delocalised behaviour in the wave description either: the highest occupied molecular orbital eigenvalue coefficients are p_z contributions (perpendicular to the plane of the molecule), but the direction of the O3 p_z-function is opposite to that of the C1-C2 and C4-C8 functions, forming a diene plus oxygen lone-pair structure. This is mirrored by the chromone wave function: the HOMO coefficients are all p_z ones, with the O3 p_z in the opposite direction from those of C1, C2, C4, C8 and C5.

The other ring, a fused benzene ring, has bond lengths varying from 1.376 to 1.403 A and is delocalised in character, even though the bond lengths C5-C6 and C9-C10 are noticeably shorter than the others.

Now that there was a chromone skeleton available, it was possible to add on the phenyl ring and the hydroxyl groups to produce the flavone compounds. It was thought that this might influence the geometry of the 3-hydroxy group, so an initial investigation of the effect of 2-substituents on the 3-hydroxy and 4-carbonyl was undertaken. The derivatives chosen were based on a range of simple groups, namely OH, NH_2, F, and Me. The carbon skeleton was kept constant and the geometry of the 2-substituent and the 3-hydroxy and 4-carbonyl groups was optimised for each molecule. The results are shown in Table 7-3. It can be seen that, apart from the dihydroxy compound, the geometry of the 3- and 4-positions is unaffected - or

only marginally affected - by the substituents. The dihydroxy compound is mostly unaffected, except for the bond angles C4-C8-O12 and C7-C8-O12, which are increased and reduced by 2°, pushing the 3-hydroxy group closer to the 2-hydroxy group. This interaction is due to the hydrogen bonding interaction between the 2-hydroxy hydrogen and the 3-hydroxy oxygen. The carbonyl oxygen at position 4 stays bent towards the 3-hydroxy group and all the bond angles are the same as the corresponding ones for the unsubstituted 3-hydroxychromone.

The 3 to 4 area of the molecule was thought to be important in the activity of the flavones. As the 3- and 4-positions of the chromone ring seem to be unaffected by the 2-substitutions described above, it seems justifiable to graft on the phenyl group to construct 3-hydroxyflavone, or flavonol, without any further optimisation.

The flavones studied include all the active inhibitors which were aglycones: myricitrin, quercitrin and rutin were too large to study using the STO-3G basis set. This leaves nine flavones and flavanones with an activity range of just under two orders of magnitude, from taxifolin, with an I_{50} concentration of 330 μmol l^{-1}, to myricetin, with one of 5 μmol l^{-1}.

The inhibitory strength of these molecules seems to be allied to certain features of their structure. In general terms, the flavones are more inhibitory than the flavanones, which could either be due to the chemical properties of the 2,3-double bond of the flavones, or just to the steric consequences of the sp^2 carbon atoms, enforcing the substituents to be in the same plane as the chromone ring. Another necessary feature seems to be hydroxylation of the phenyl ring, as the inhibition increases with increasing numbers of hydroxyl groups on this ring. There is an anomaly, however, in that 3-hydroxyflavone, which

has no hydroxyl group on the phenyl ring, is as potent an inhibitor as quercetin. 3-Hydroxyflavone has no hydroxy groups on the 5- and 7-positions, either, but these do not seem to be important for activity anyway, as fisetin lacks the 5-hydroxy group too, and yet is as strong an inhibitor as quercetin. The pattern of the hydroxyl groups on the phenyl ring seems to be important, however, as morin, which is an isomer of quercetin, but has hydroxy groups at the 2'- and 4' positions, instead of 3' and 4', has an I_{50} concentration of 30 μmol l^{-1}, as opposed to 9 μmol l^{-1} for quercetin.

The above observations do not lead to a logical ordering of the inhibitors by inspection of their structures alone. It was thought that the electrostatic properties of these compounds might lead to an index of inhibitory power which could be applied to prospective inhibitor molecules. To this end, the molecular electrostatic potentials of these molecules were studied in the plane of the chromone ring and on their van der Waals surfaces. The geometries of the flavone and flavanone derivatives were based on the 3-hydroxychromone structure. The phenyl ring was constructed from standard bond lengths and angles. Many discussions of the flavones seem to imply that these compounds are completely planar, but it is immediately apparent to anyone who has built a CPK model that the substituents on the 2'- or 6'-positions are obstructed by the 3-substituent, preventing a planar structure, or free rotation. It was discovered from calculations using the PCILO method [177] on 3-hydroxyflavone that the phenyl ring's preferred position, upon rotation around the phenyl-carbon bond, is at an angle of 50° to the plane of the chromone ring. This 50° dihedral angle also allows the hydrogen atom at either 2' or 6' (depending which way the phenyl ring is tilted) to interact with a lone pair of the 3-hydroxy oxygen.

It was felt that it would be worth finding out if the MEP on the van der Waals surface of rings A and C of 3-hydroxyflavone were affected by the rotation of the phenyl ring. Therefore three orientations of the ring were used in the MEP calculations. These conformations involved placing the ring at right angles to the chromone portion of the molecule, at 50° to it and in the same plane as it. As can be seen from Photographs 5 to 7, the chromone ring is unaffected at the van der Waals distance for a 90° to 50° dihedral angle, but the planar molecule does have a different electrostatic potential on the 3-hydroxy oxygen; this is not too surprising, as the 2'-hydrogen is touching this oxygen, being only 1.47 A away from it. As the chromone part of the molecule did not seem to be particularly affected by reasonable orientations of the phenyl ring, it was decided to calculate the MEP of the other flavone and flavanone compounds using just the 50° dihedral orientation. As this gives an asymmetric molecule it was thought necessary to calculate the MEP for all the possible isomers of the molecules which fulfilled the following criteria: the hydroxy groups on the phenyl ring B had to planar to this ring and they all had to "point" the same way around the ring; that is, the hydrogen of one points towards the oxygen of the next one. This gives two calculations for myricetin (Photographs 8 and 9), four for quercetin (Photographs 10 to 13), four for fisetin (Photographs 14 to 17), four for morin (Photographs 18 to 21), two for kaempferol (Photographs 22 and 23) and two for apigenin (Photographs 24, one isomer not shown). Only single isomers of the flavanone molecules were used in the calculations, as it was felt that the lack of activity was more to do with the gross steric configurations of these molecules, rather than more subtle electrostatic effects. Naringenin is shown in Photograph 25 and taxifolin in Photograph 26.

The most inhibitory flavone is myricetin, with an I_{50} concentration of 5 μmol l^{-1}. This is 3,5,7,3',4',5'-hexahydroxyflavone, having a triply-substituted phenyl ring, like pyrogallol. The MEP of the two isomers obtained by rotating either the hydroxy groups or the phenyl ring by 180° (it amounts to the same thing) are mostly similar except on the surface of the phenyl ring (as might be expected). The rotation of the hydroxy groups does influence the potential on atoms as far away as the 8-position of the chromone ring, but it is only a slight effect, which is within the range of error associated with an electrostatic potential calculation using a minimal basis set. The main effect is noticeable on the phenyl ring in the neighbourhood of the 3-hydroxy group; depending on the orientation of the hydroxy groups, there is a negative region running from the 3-hydroxy oxygen to the base of the 3'-hydroxy group, or the 5'-hydroxy group, depending on which one is pointing away from the 3-hydroxy group. The actual potential on the 2'- or the 6'-carbon, whichever is nearest to the 3-hydroxy group, is very negative, from the influence of the oxygen lone pairs.

The next most inhibitory flavones are 3-hydroxyflavone, itself, and quercetin - 3,5,7,3',4'-pentahydroxyflavone - which both have an I_{50} concentration of 9 μmol l^{-1}. Fisetin - 3,7,3',4'-tetrahydroxyflavone - with an I_{50} concentration of 10 μmol l^{-1}, is of comparable strength, too. As mentioned above, the apparently anomalously high value of 3-hydroxyflavone is the interesting feature of this group of compounds. If we ignore this molecule, a rationale of the inhibitor strength on the basis solely of the hydroxylation of the phenyl ring and the degree of saturation of the 2,3-bond would be possible. However, this breaks down when applied to 3-hydroxyflavone, as it has no phenolic groups. The region of

negative potential noted in the myricetin calculations is also present on the 3-hydroxyflavone surface at this point. It is more extensive than the myricetin negative region, but it is not as directional, forming an almost symmetrical negative region on the phenyl ring surface.

The negative region of the other two species is comparable for corresponding pairs of rotamers (i.e. Photographs 10 and 14, 11 and 15, 12 and 16, and 13 and 17). These do not have as extensive a negative region as the 3-hydroxyflavone, but have directed paths of negative potential leading to the base of the nearest hydroxy group pointing away from the 3-position, as was the case with the myricetin molecules. These negative regions are not as narrow as the myricetin regions, though.

The morin - 3,5,7,2',4'-pentahydroxyflavone - molecules differ from the others in that they have a 2'-hydroxy group. Morin's I_{50} concentration is slightly less than its isomer, quercetin, being 30 μmol l^{-1}. The presence of the 2'-hydroxy group on the neighbouring chromone ring is much more profound than for the other molecules; the EP of the pyrone part of the system changes quite dramatically, depending on whether the hydroxy group points towards or away from it. As this region does not appear to vary much for the other molecules it is not diagnostic of inhibitors' efficacy, so we shall concentrate on the phenyl ring instead. This again has a negative region which is most negative near the 3-hydroxy group. This area extends directly across the ring to the 4'-position, not to the side as in the case of myricetin. The same is true of kaempferol - 3,5,7,4'-tetrahydroxyflavone - this has an I_{50} concentration of 55 μmol l^{-1}. There is a small, strongly negative region near the 3-hydroxy group and a less negative region extending to the base of the

4'-hydroxy group.

Apigenin - 5,7,4'-trihydroxyflavone - lacks the 3-hydroxy group present in the other molecules. This has a very different electrostatic potential on the van der Waals surface of the phenyl ring, being mostly slightly positive. It still possesses the 4-carbonyl group and the unsaturated bond between the 2- and 3-positions.

A rationale for the inhibition of the enzyme by these compounds can now be put forward. Assuming that the first event in the enzyme-catalysed reaction is the binding of the glutathione part of the substrate to a binding site in the enzyme, it would still be necessary to have some sort of interaction to push the hemithioacetal part towards the catalytic mechanism, as glutathione is quite a flexible molecule [178]. This could be achieved by placing a positive charge (or positive region) where the sulphur atom was to be. The more negative oxygen atoms (initially a hydroxy and a carbonyl group) would be held in position by the zinc atom via a bound water molecule as shown in Figure 3-1. The proton-abstracting base could then remove the hydrogen and donate it to the other carbon atom.

The 3-hydroxyflavone compounds possess the same CO-COH moiety as the substrate and also have a negative region on the phenyl ring where the sulphur atom would be in the substrate. Thus they could bind quite effectively to the oxygen and sulphur binding sites. However, it is not possible for the abstracting base to remove a proton as before, thus incapacitating the enzyme until the flavone dissociates form the active site.

Both of the myricetin rotamers have a strong local negative region on the phenyl ring surface, especially the conformation shown in Photograph 9. Photograph 8 shows that the other conformer has this region plus a narrow area of negative potential extending to the base of the 3'-hydroxy group. The rest of the phenyl surface is slightly positive. The next strongest inhibitors are quercetin and fisetin, which can be considered together, and 3-hydroxyflavone. Quercetin and fisetin show the same negative region, but it is not as narrowly defined on the surface as the myricetin potential, extending over most of the phenyl surface. 3-Hydroxyflavone has even less directional shape to the negative region, but it is more negative over a larger area, perhaps compensating for this. Morin has the same sort of pattern as quercetin and fisetin, but it is distorted by the proximity of the 2'-hydroxy group, which either diminishes the depth of the negative region, as on Photographs 18 and 21, or pulls it sideways, as on Photographs 19 and 20. As mentioned before, the 3- and 4-positions have differing potentials on their surface, the conformer of Photograph 21 resembling the other flavones the most. All these factors would tend to reduce the interaction with a positive site, or at least translate it sideways, so that the 3-hydroxy,4-carbonyl groups are not in exactly the right place for a strong interaction with the zinc atom. Kaempferol has a small negative region of less than -0.015 electrostatic units and a larger almost symmetrical region of negative potential, which again does not cover just the 2' to 6' portion of the surface as it does for myricetin. Apigenin could only interact with the zinc atom through the 4-carbonyl group, as it has no 3-hydroxyl group and hence no negative region on the phenyl surface.

The MEP on the van der Waals surface of two other flavonoids was also calculated. These were luteolin, 5,7,3',4',5'-pentahydroxyflavone and robinetin, 3,7,3',4',5'-pentahydroxyflavone. These are shown in Photographs 27 and 28. Luteolin does not have the required negative region on the phenyl ring and so is predicted to be of only moderate activity, as it has the 4-carbonyl group and the 2,3-double bond. Robinetin, on the other hand has a similar pattern to myricetin and would be expected to be quite an active inhibitor.

7.2 The Coumarin Compounds

Subsequent to the discovery of the activity of the flavone compounds, it was found that a range of hydroxylated coumarin derivatives were also active [107]. In fact, one of them was found to be more inhibitory than myricetin itself. These coumarins are shown in Figure 7-2.

The coumarin group of molecules are related to chromones in that they are benzopyrones, but they have an α-pyrone ring instead of a γ-pyrone ring. Like the flavonoids, they are found in plants as secondary metabolites, particularly in the rutaceae and umbelliferae families [179]. They have wide-ranging biological actions, notably the aflatoxins, 4-hydroxycoumarins and furanocoumarins. The aflatoxins, which are complex molecules with a fused ring structure and a central coumarin skeleton, are carcinogenic and are particularly associated with liver cancers. Some derivatives of 4-hydroxycoumarins are anti-coagulants, whilst others are antibiotics. The furanocoumarins have a wide range of activities based on the photobinding reaction in

which these compounds bind to protein and nucleic acids upon irradiation by ultraviolet light [179].

Of more interest in this work are reported incidences of tumour inhibition and enzyme inhibition. Esculetin - 6,7-dihydroxycoumarin - is known to inhibit rat platelet lipoxygenase and cyclooxygenase activities [180]. The investigators also found that esculin, which is the 6-glucoside of esculetin, and umbelliferone - 7-hydroxycoumarin - also inhibited lipoxygenase, but 4-hydroxycoumarin and coumarin itself did not inhibit it. It has also been found that esculetin inhibits tumour development in mouse skin, initiated by application of benzo[a]pyrene [181]. Ohta et al. [182] found that coumarin and umbelliferone possessed antimutagenic activity against ultraviolet irradiation of E. coli, similar to the previously reported inhibition by cinnamaldehyde [183].

Some naturally occurring coumarins have also been found to induce some drug-metabolising enzymes found in the liver, which metabolise foreign compounds as part of the detoxication mechanism [184].

The coumarin derivatives investigated in this work are hydroxylated versions, some of which were synthesised for the first time in the course of the studies [107]. The activity of these hydroxycoumarins in inhibiting glyoxalase I catalysis was wider in range than the flavonoid series, varying from an I_{50} concentration of 3.5 μmol l^{-1}, for coumarin No. 8, the most active compound so far discovered, to 1400 μmol l^{-1} for coumarin itself. It is not as easy to order these compounds based on structural relationship to their activities, as it was for the flavones, so it was thought that the MEP calculations would provide more useful information.

The coumarin structure was optimised using the Murtaugh-Sargent optimisation procedure [133] in GAUSSIAN 82 [130]. This proved to be very difficult, as the structure was quite rigid, producing large changes in the forces on the atoms for small changes in geometry. The optimisation was achieved by restricting the step size allowed to the optimisation routine to a quarter of its usual value. This (eventually) resulted in the optimised geometry listed in Table 7-4 and shown in Figure 7-4.

The structure of the two rings, as is the case of the 3-hydroxychromone calculation, shows an aromatic benzene fused to an α-pyrone ring which is not delocalised: the geometry is similar to that of the 3-hydroxychromone molecule, except that the bond lengths between the 2- and 3- and the 3- and 4-positions are swapped around, with the transfer of the carbonyl oxygen atom from the 4-position to the 2-position. This also increases the length of the O1-C2 bond. The fused benzene ring shows the same slightly shortened bonds between positions 5 and 6 and 7 and 8, as for the 3-hydroxychromone.

The studies on the derivatives of this molecule were carried out using the coumarin skeleton and limited optimisation of the substituent groups. The MEP on the van der Waals surface of these compounds were calculated and are shown in Photographs 29 to 39.

The strongest inhibitor is coumarin No. 8 - 7,8-dihydroxy-4-phenylcoumarin - which has an I_{50} concentration of 3.5 μmol l^{-1}. Most of the rest of the compounds tested have hydroxy or methoxy groups in the 6- and 7-positions, except for isoesculetin - 3,4-dihydroxycoumarin - and esculin, which is the glycoside of esculetin (6,7-dihydroxycoumarin).

The next strongest inhibitors after coumarin No. 8 are coumarins No. 7 and No. 10, which are 6,7-dihydroxy-4-(4-fluorophenyl)coumarin and 6,7-dihydroxy-4-phenylcoumarin respectively. These both have I_{50} concentrations of 20 μmol l^{-1}. If the 4-phenyl group of coumarin No. 10 is replaced by a benzyl group, then the molecule resulting, coumarin No. 4, is slightly less active, having an I_{50} concentration of 24 μmol l^{-1}. Coumarin No. 5 - 3-benzyl-6,7-dihydroxy-4-methylcoumarin - has an I_{50} concentration of 30 μmol l^{-1} and differs from No. 4 in that the benzyl group has been moved to the 3-position and the 4-position has a methyl group. When the 3-benzyl group is removed the molecule is 4-methylesculetin (6,7-dihydroxy-4-methylcoumarin) which is less active again, with an I_{50} concentration of 50 μmol l^{-1}. Compound No. 9 is similar to this apart from the fact that the methyl group has become a propyl group, with a corresponding loss of inhibitory power, down to 80 μmol l^{-1}.

Isoesculetin, which contains the nearest structure found in these compounds to an enediol structure has an I_{50} concentration of only 145 μmol l^{-1}. Its is known as 3,4-dihydroxycoumarin.

Methoxy derivatives of coumarin No. 10, coumarins No. 1 and No. 2, display a great loss of activity, with I_{50} concentrations of 150 and 200 μmol l^{-1} and esculetin - 6,7-dihydroxycoumarin - which does not possess the 4-phenyl group is not very active, either, with an I_{50} concentration of 200 μmol l^{-1} too.

The least active compound in the series is coumarin No. 3 - 6,7-dihydroxy-3,4-dimethylcoumarin. This has an I_{50} concentration of just 320 μmol l^{-1}, and yet it only differs from 4-methylesculetin by one methyl group.

The van der Waals surface MEP of these compounds (Photographs 29 to 39) are more complicated than those of the flavonoid series. The general form of this surface is to have a very negative region on the ring oxygen and carbonyl oxygen and a negative region on the hydroxyl oxygen of the hydroxyl which has its hydrogen pointing towards the other hydroxyl oxygen. The negative region on the second hydroxyl group is reduced by the influence of the positive hydroxyl hydrogen to which it is hydrogen bonded. The rest of the coumarin fused rings' surface is slightly positive, apart from small, thin regions of negative potential extending from the oxygen atoms towards the middle of the molecule. The substituent phenyl rings have central portions of negative potential similar to the 3-hydroxyflavone phenyl ring, but they are not as large as for the flavone.

Coumarin No. 8, the most active species, has a slightly different pattern of negative potential on its surface from the less active molecules: the negative region extends from the base of the 7-hydroxyl oxygen instead of the 6-hydroxyl oxygen as in the other molecules, and the positioning of the negative and positive regions associated with the hydroxyl groups has shifted round to the 7- and 8-positions, from the 6- and 7-positions.

The next two most active species are coumarins No. 7 and No. 10. Photograph 29 shows the van der Waals electrostatic potential surface of coumarin No. 10. there is a small difference in the electrostatic potential of this from that of No. 7, in that the MEP for molecule No. 7 does not have a negative region on the tip of its phenyl ring, but has a substantial negative potential surface on the fluorine atom substituent of this ring. The pattern of the rest of the surface is essentially the same.

Coumarin No. 4, which is slightly less active than the two species above, has a similar pattern of potential on its coumarin skeleton surface, but the CH_2 group of the benzyl substituent forces the phenyl group to occupy a different orientation, leaning away from the perpendicular. Thus, this species may not fit in the active site of the enzyme as well as molecules 7 and 10. Coumarin No. 4 is shown in Photograph 30.

The next four photographs (Photographs 31 to 34) show different conformers of coumarin No. 5, which has the benzyl ring at position 3. It was decided to investigate different orientations of the benzyl group, as it seemed possible, from model building studies, that the phenyl part of the ring could occupy a region of space which was quite close to that occupied by the phenyl groups of molecules 7 and 10. The MEP on the coumarin skeleton near the hydroxy groups and the carbonyl group does not change much unless the benzyl group is positioned adjacent to the methyl group, as in Photograph 33. This is the orientation mentioned above, but it turns out to be the least stable of the conformers energetically. However, the electrostatic potential on the phenyl ring is influenced by the proximity of the carbonyl group, similar to the 3-hydroxyflavone case. There is still a region of negative potential running from the base of the 6-hydroxy group into the centre of the ring system, but it is not as prominent as for the more powerful inhibitors.

4-Methylesculetin (Photograph 35) has a similar electrostatic potential to coumarins No. 7 and No. 10, but lacks the hydrophobic phenyl group at position 4, which is replaced by the smaller methyl group. There is very little difference between coumarin No. 10 and this compound apart from a small region of negative potential on the

3-carbon.

Coumarin No. 9 is less active than apigenin, but differs from the more active coumarins only in having a propyl group at position 4, which is slightly positive on its van der Waals surface, and a less negative region at position 6. It seems from an examination of the van der Waals surface near carbon 6, that this region may be of use in indicating the strength of the inhibition of the coumarins: the larger the negative portion of this surface is, the more inhibitory the molecule is. Isoesculetin (Photograph 36) has an unsubstituted fused benzene ring, unlike the other coumarins, and the pattern on this ring is very different from the others, forming a curved negative region on the inside of the carbon atoms. There is no negative region extending from the base of the hydroxy oxygens on the ∝-pyrone ring part, either.

The other exception to the general series of strength of the inhibitors, apart from isoesculetin, is coumarin No. 3 (Photograph 38), which is not very inhibitory at all. This has a similar pattern of potential to coumarin No. 10 on the benzo ring, but the negative potential near the carbonyl group is extended over the surface of the 3-methyl group. This extension of the negative region is not present on the surface of any of the other molecules.

The other poor inhibitor is the unsubstituted coumarin (Photograph 39) which is not inhibitory to any real extent. This also has an area of potential which is not present on any of the other molecules: the 4-position, which has a substituent in the other molecules, has a region of potential which is greater than 0.030 a.u. on the tips of the 4- and 5-hydrogens.

The above results, tend to show that, although the MEP on the van der Waals surface of these molecules does not appear to be useful as an index of inhibitory power, as it proved to be for the flavonoid compounds, there do seem to be features of it which hint at such a relationship. Therefore it was decided to investigate the electrostatic potential of these compounds further. The MEP of the coumarins was calculated at double the van der Waals radius of the atoms. At this distance the MEP is probably a more accurate representation of the interaction between a positive point charge and the molecule, as the perturbation of the structure will not be as great. The photographs of these calculations are numbers 40 to 48. It turns out that the resultant pictures are indeed easier to order in terms of activity than the previous van der Waals radius pictures.

Coumarin No. 8 (Photograph 40) has a negative region around the 1- and 2-positions, where the O-CO part of the molecule is, and a positive region encompassing the rest of the molecule. This is most positive to the left of the picture, where the 6-position is. The next most active species, coumarins No. 7 and No. 10, are almost identical to each other on this surface, apart from at the top of the phenyl ring, where the No. 7 species has a small area of negative potential (Photograph 41). The large negative region extends further over the top of the molecule in the pictures, to the vicinity of the 8-position, and the positive region moves further round towards the 5-position. Coumarin No. 4 (Photograph 42) has an even larger negative region, extending further over the top of the picture, and the positive region has moved a little further towards the benzyl ring. This trend continues down the series: coumarin No. 9 has a larger negative area extending over half the coumarin surface (Photograph 44) and coumarins No. 1 and No. 2 (Photographs 45 and 46) have the same region, but with two areas

contained in it of less than -0.015 a.u. Coumarin No. 2 also has a distinctly more positive region than No. 1. The least inhibitory molecule, coumarin No. 3 has the largest negative region, which spans from the 7-position across to the 3-position. As noted for the van der Waals surface calculation, this negative area is not so obvious on any other molecule.

The following features can now be suggested as indicators of the inhibitory power of the hydroxylated coumarins, based on the MEP results:

The molecule must possess adjacent hydroxy groups on the benzo ring of the coumarin. A large hydrophobic group, such as a phenyl, at the 4-position is also important. The consequence of this is to limit the extent of the positive potential around position 5. If the molecule has concentrated regions of negative and positive potential at the 2- and 5-positions, or more preferably, at the 2- and 6-positions, then it will be a good inhibitor. The more spread out either of these regions is, the less specific the interaction with the positively and negatively charged parts of the active site, and the less strong the subsequent association.

Table 7-1 Inhibitors Of Glyoxalase I And Their Corresponding I_{50} Values, In μmol l^{-1}.

COMPOUND	I-50
Coumarin deriv. 8	3.5
Myricetin	5
Quercetin	9
3-Hydroxyflavone	9
(Purpurogallin	9)
Fisetin	10
Coumarin deriv. 7	20
Coumarin deriv. 10	20
Coumarin deriv. 4	24
Coumarin deriv. 5	30
Morin	30
4-Methylesculetin	30
Esculetin	30
Kaempferol	55
(Quercitrin	70)
Apigenin	70
Pyrogallol	70
(Pyrogallol Red	70)
Coumarin deriv. 9	80
(Rutin	110)
Coumarin deriv. 1	150
(Myricitrin	185)
Coumarin deriv. 2	200
Naringenin	250
(Esculin	270)
Coumarin deriv. 3	320
Taxifolin	330
Coumarin	1400

The compounds in brackets have not been studied by this group.

Table 7-2 The STO-3G Optimised Geometry Of 3-Hydroxychromone, In Angstroms And Degrees.

C1-C2	1.395
C1-O3	1.392
O3-C4	1.392
C1-C5	1.403
C5-C6	1.376
C2-C7	1.497
C4-C8	1.325
C7-C8	1.504
C2-C9	1.399
C6-C10	1.397
C9-C10	1.376
C7-O11	1.228
C8-O12	1.396
O12-H13	0.992
C-H	1.083
C2-C1-O3	124.2
C1-O3-C4	116.4
C2-C1-C5	120.2
O3-C1-C5	115.6
C1-C5-C6	119.6
C1-C2-C7	119.4
C4-C8-C7	121.1
O3-C4-C8	125.5
C2-C7-C8	113.5
C1-C2-C9	119.2
C7-C2-C9	121.4
C6-C10-C9	119.9
C5-C6-C10	120.6
C2-C9-C10	120.6
C2-C7-O11	126.1
C8-C7-O11	120.4
C4-C8-O12	122.2
C7-C8-O12	116.7
C8-O12-H13	102.6
O3-C4-H14	111.7
C-C-H	119.3

Table 7-3 The STO-3G Optimised Values Of Certain Bond Lengths Of Various 2-Substituted 3-Hydroxychromones, In Angstroms And Degrees.

	2-Substituent				
	None	OH	Me	NH_2	F
C7-O11	1.228	1.230	1.230	1.230	1.229
C8-O12	1.396	1.405	1.398	1.397	1.399
O12-H13	0.992	0.990	0.992	0.992	0.992
C2-C7-O11	126.1	125.0	125.2	125.3	125.5
C8-C7-O11	120.4	120.4	120.1	120.1	119.8
C7-C8-O12	116.7	118.9	116.8	117.5	117.1
C4-C8-O12	122.2	120.4	122.5	121.8	122.3
C8-O12-H13	102.6	102.6	102.3	102.4	102.2

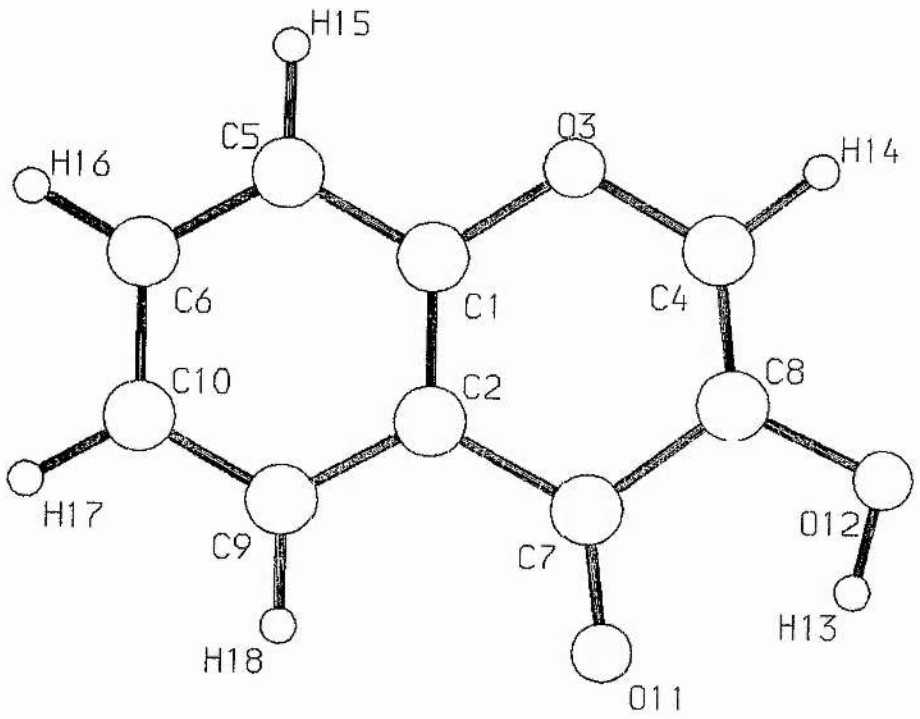
Table 7-4 The STO-3G Optimised Geometry Of Coumarin, In Angstroms And Degrees.

C1-C2	1.397
C1-O3	1.391
O3-C4	1.407
C1-C5	1.400
C5-C6	1.378
C2-C7	1.474
C4-C8	1.500
C7-C8	1.322
C2-C9	1.398
C6-C10	1.396
C9-C10	1.398
C4-O13	1.219
C-H	1.083
C2-C1-O3	123.7
C1-O3-C4	119.0
C2-C1-C5	120.4
O3-C1-C5	115.8
C1-C5-C6	119.5
C1-C2-C7	117.8
C4-C8-C7	122.3
O3-C4-C8	117.5
C2-C7-C8	119.6
C1-C2-C9	119.0
C7-C2-C9	123.2
C6-C10-C9	120.0
C5-C6-C10	120.5
C2-C9-C10	120.5
O3-C4-O13	117.2
C8-C4-O13	125.3
C-C-H	120.0

Figure 7-1 The Flavonoid Compounds.

Figure 7-2 The Coumarin Compounds.

Figure 7-3 The STO-3G Optimised Geometry Of 3-Hydroxychromone.



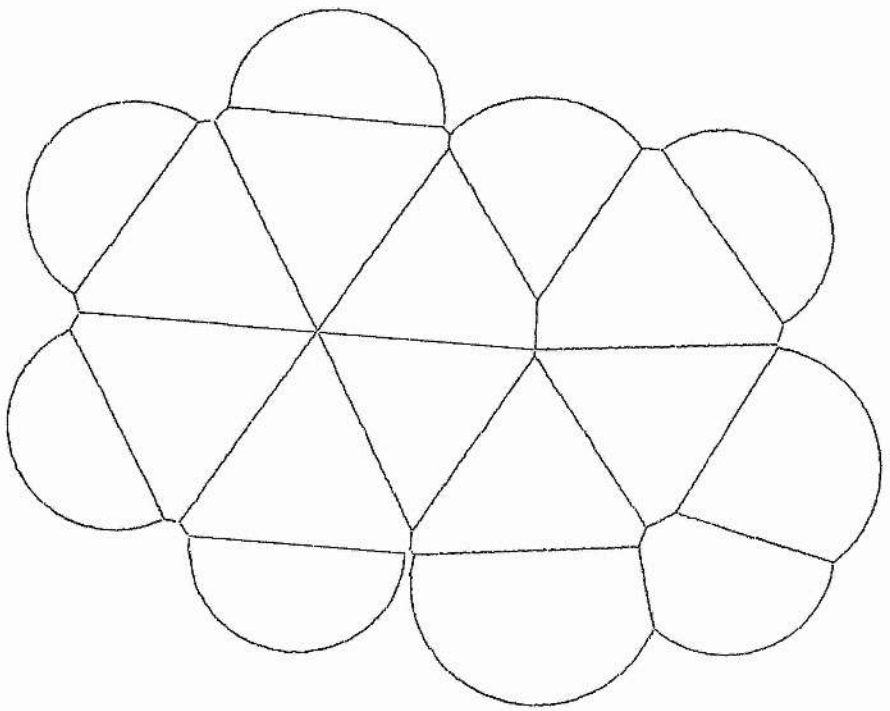
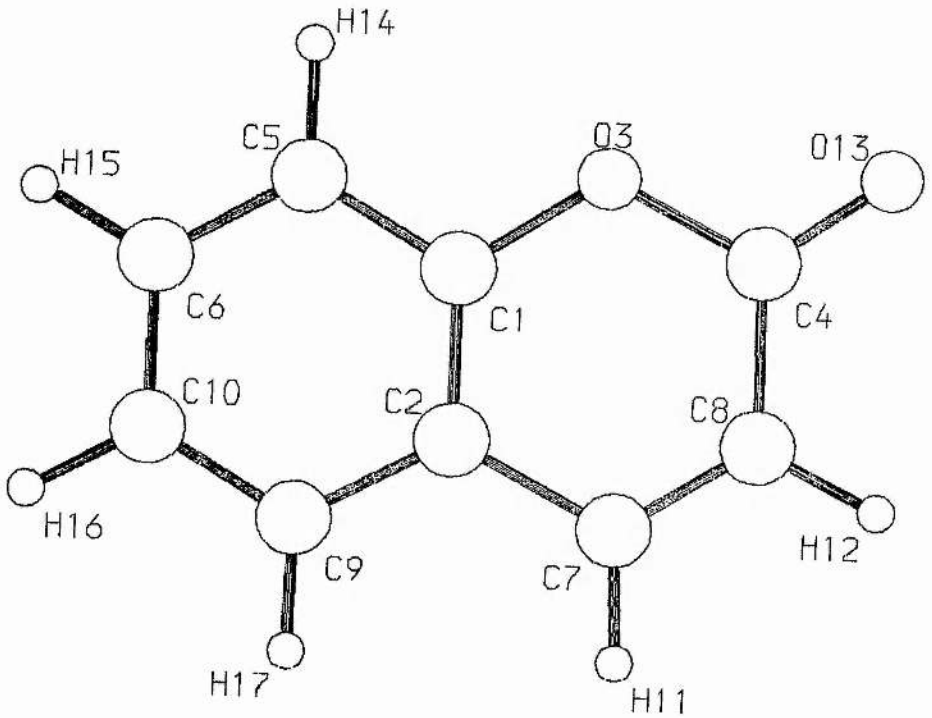
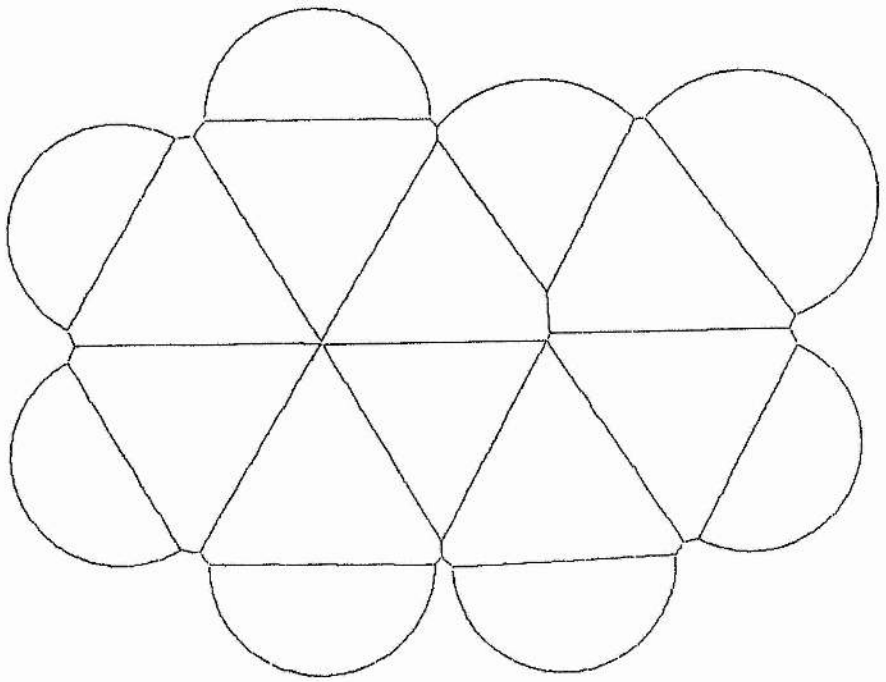


Figure 7-4 The STO-3G Optimised Geometry Of Coumarin.





CONCLUSIONS

The results discussed in the previous chapters are summarised below. It should be noted that none of the work is definitive, in that the basis sets used are far from complete, and no account has been taken of correlation effects. Nevertheless, it is felt that the results provide a stimulus for further work on the design of enzyme inhibitors and on the elucidation of the enzyme mechanism.

The model reaction scheme results support the latest work on the enzyme mechanism [89,90], in predicting enediol intermediates in the enzyme reaction. The route through the enediol intermediate by alternate deprotonation and protonation reactions is the most favourable energetically in all the reaction schemes studied. These included calculations in the absence and presence of magnesium, with minimal and split-valence basis sets.

The comparison of fully optimised geometry calculations using the 3-21G and 3-21G* basis sets demonstrates that the sulphur d-functions do affect the structure of the molecules noticeably, particularly that of molecule F. The comparison of these optimised structures with those of molecules in which the thiol group was replaced by a hydroxyl group or a hydrogen show that the sulphur atom stabilises the anions relative to the neutral molecules, reducing the energy barriers in the proposed reaction.

The introduction of a magnesium 2+ ion, to represent the zinc atom at a crude level, results in the "stabilisation" of the previously unstable anionic compounds, molecules B, D and G. However, none of the wave functions for these magnesium adducts had the correct number of bound molecular orbitals, so some caution must be exercised in drawing

conclusions from these results. The magnesium ion inclusion reduces the energy barriers substantially: down to a level where imidazole could catalyse the reaction. This base is part of histidine, which is believed to be part of the active site [69]. This particular scheme bears a direct relationship to the model scheme of Hall et al. [81,82].

A justification for the use of such small species as models for the large hemithioacetal-derived molecules is provided by the further work in which the thiol hydrogen was replaced by methyl and ethyl groups. The resultant optimised geometries are remarkably similar to the parent molecules' geometries, as are the results from population analysis studies of the corresponding wave functions.

The FORTRAN graphics program which was developed in this work has been in use by Dr Colin Thomson's research group for over a year now. It is being used in other projects studying enzyme inhibition and tumour promotion. The speed and flexibility of the interactive features of the program are good, and the resulting representations are easier to interpret than the more traditional chicken-wire mesh. The modifications suggested at the end of Chapter 5 would turn the program into a more flexible version, in which the user could change the contouring of the display interactively and also directly compare molecules of varying sizes, by specifying an absolute scale. These modifications would require little extra work.

The work on the inhibitor molecules is based heavily on the graphics program. It was discovered that certain regions of the inhibitor molecules' MEP surface, either at a van der Waals distance, or at double this distance, correlated well with inhibitory strength.

The flavone compounds were constructed using an optimised chromone geometry, at the RHF/STO-3G level. It was found that the substituent at the 2-position did not affect the geometry of the 3-hydroxy group much. This justified the use of single-point calculations of the flavone molecules, instead of prohibitively expensive optimisations, even of small parts of the structure, such as the hydroxy groups. It was found that the orientation of the hydroxyl groups affects the pattern of the MEP on the phenyl ring surface, so different orientations of these hydroxyl groups were used in the calculations.

The only differences of the MEP on the van der Waals surface of the flavones which seemed to correlate with the inhibitor strength was the pattern of the potential on the phenyl ring. It was suggested that this area of the molecule is in contact with a part of the enzyme responsible for binding the sulphur atom of the substrate.

The hydroxylated coumarin compounds were constructed from the geometry of coumarin itself, which had been optimised using the STO-3G basis set. No good correlation between the MEP on the coumarin derivatives' surfaces and their inhibitory power was noted. However, it was decided that plots of double the van der Waals surface might be of more use. This proved to be the case, as a correlation between the positioning of a negative region of potential near the 1- and 2-positions and a positive region near the 5- and 6-positions was found.

The next step in this work could be to consolidate the correlative data mentioned above by comparing the molecules to each other on a standard surface, which would allow a direct comparison of the electrostatic potential at certain points in space. This is analogous

to difference maps obtained from planar slices through similar molecules. There are two sensible ways that this can be done: The surface of the largest species could be used, or the surface of the most active species. The former method allows the comparison of all the molecules without the danger of the surface passing through any atomic nucleus, giving huge positive potential values, which can obviously not be compared with potential values from the surface of other molecules. The latter method has an attractive feature if we assume that the most active species fits into the enzyme or receptor site better than the other molecules. An examination of the potential on the van der Waals, or double van der Waals surface which fits best into the site may give information not only on why the other molecules do not interact as strongly, but also on the nature of the site itself.

The model reaction scheme studies can be extended by investigating the use of other basis sets, notably ones with diffuse functions. In my opinion it would be found that the anionic species are stabilised further in comparison with the neutral species. The interaction of the bound water which is apparently attached to the zinc atom [67,68,69] with the substrate could be investigated by carrying out supermolecule calculations; firstly using water and protonated water species, but maybe with small magnesium compounds too. Some work has been done involving the interaction of water and H_3O^+ with members of the basic reaction scheme, but this work is not complete.

The renewed interest in the glyoxalase I enzyme stemmed not only from Szent-Gyorgyi's proposals concerning its rôle in tumorigenesis [14,44,56,57], but also from the controversy surrounding its supposed mechanism, which was originally thought to be analogous to the Cannizzaro reaction [76-78]. Some of the model reactions investigated

to shed some light on this enzyme were indeed found to proceed via rate-limiting hydride-transfer reactions [79-80]. It would be interesting to compare model hydride-transfer reaction schemes with the proton transfer schemes reported here.

INDEX OF PHOTOGRAPHS

1) Molecule A, sulphur side.
2) Molecule A, hydrogen side.
3) Molecule E
4) Molecule H
5) 3-Hydroxyflavone, phenyl ring at right angles.
6) 3-Hydroxyflavone, phenyl ring at 50 degrees.
7) 3-Hydroxyflavone, phenyl ring parallel.
8) Myricetin, rotamer No 1.
9) Myricetin, rotamer No 2.
10) Quercetin, rotamer No 1.
11) Quercetin, rotamer No 2.
12) Quercetin, rotamer No 3.
13) Quercetin, rotamer No 4.
14) Fisetin, rotamer No 1.
15) Fisetin, rotamer No 2.
16) Fisetin, rotamer No 3.
17) Fisetin, rotamer No 4.
18) Morin, rotamer No 1.
19) Morin, rotamer No 2.
20) Morin, rotamer No 3.
21) Morin, rotamer No 4.
22) Kaempferol, rotamer No 1.
23) Kaempferol, rotamer No 2.
24) Apigenin, rotamer No 1.
25) Naringenin.
26) Taxifolin.
27) Luteolin.
28) Robinetin.
29) Coumarin No 10.
30) Coumarin No 4.
31) Coumarin No 5, rotamer No 1.
32) Coumarin No 5, rotamer No 2.
33) Coumarin No 5, rotamer No 3.
34) Coumarin No 5, rotamer No 4.
35) 4-Methylesculetin.
36) Isoesculetin.
37) Coumarin No 1.
38) Coumarin No 3.
39) Coumarin.
40) Coumarin No 8, double van der Waals surface.
41) Coumarin No 7, double van der Waals surface.
42) Coumarin No 4, double van der Waals surface.
43) 4-Methylesculetin, double van der Waals surface.
44) Coumarin No 9, double van der Waals surface.
45) Coumarin No 1, double van der Waals surface.
46) Coumarin No 2, double van der Waals surface.
47) Esculetin, double van der Waals surface.
48) Coumarin No 3, double van der Waals surface.

APPENDIX

GRAPHICS PROGRAM LISTING

This program listing is the VAX version of the graphics program. The Tektronix DTI routines have been converted to equivalent versions on the VAX using terminal escape sequences. These graphics primitive routines all start with the letters CE, as opposed to the DTI versions which have the initial letters LL. The graphics routines are stored in a library, which is linked with the program object module at compile time. They are listed after the program, along with the required encoding and decoding routines.

```
CE  CE3D1.FOR :  a program for drawing a rough representation
CE of a 3-dimensional object coded as coloured squares
CE with indices from 0 to 15
CE It builds up the picture from the lowest to the highest
CE z-coordinate
CE It gives an impression of depth by angling the squares
CE according to the vector normal to the surface at that point
CE It draws only those squares facing towards the observer
CE This version has an improved sorting routine.
CE
CE                Colin Edge, September 1985
CE     (based on a program from Andy Micklethwaite, Tektronix)
CE
CE    Initialise
       COMMON /TRANSFORM/SCL(4,4),RX(4,4),RY(4,4),RZ(4,4),
      + UNIT(4,4),T(4,4),W(4,4)
       COMMON /WINDAT/XLEFT,XRIGHT,YBOTTOM,YTOP,XSCALE,YSCALE,
      + XDIFF,YDIFF
       COMMON /MENU/INFILE,CORIENT,CD,CS,CLIGHT,CVIEW,CDIAG,
      + CISEG,CIVIEW
       COMMON /VALUES/ ROTX,ROTY,ROTZ,D,S,XL,YL,ZL,XE,YE,DIAG,
      + ISEG,IVIEW,OCTAGON,BACKING,ROTATION
       CHARACTER*15 INFILE,CFILE
       REAL COORD(3,4000),TRANS(3,4000),TRANS2(3,4000)
       REAL PERP(3,4000),PERPT(3,4000)
       REAL D,S
       REAL XE,YE
       REAL XL,YL,ZL
       REAL SCALE,ROTX,ROTY,ROTZ,DIAG
       LOGICAL ORIGIN,OCTAGON,BACKING,NOVICE,FAST,ROTATION,TIMER
       INTEGER ICOL(4000),ICOLT(4000)
       INTEGER IPX,IPY
       INTEGER NCOUNT,NT,IONE,ISEG,IVIEW,I14
       CHARACTER*6 CDIAG,CROTX,CROTY,CROTZ
       CHARACTER*6 CXE,CYE,CXL,CYL,CZL,CD,CS
       CHARACTER*5 CISEG,CNCOUNT,CIVIEW
       CHARACTER*18 CORIENT,CLIGHT
       CHARACTER*12 CVIEW
       CHARACTER*1 CANS
       CALL CEINIT(ITERM)
CE Set 'S' and '' as report signature characters
       CALL CERPSG(0,83,0)
       CALL CEDACL
       WRITE(6,101)'            3D Design Program'
 101   FORMAT(A,//)
       CALL WINDOW(0.,0.,100.,100.)
CE   Read in coordinates
       IPX=0
       IPY=0
       I14=-14
       SCALE=1.
       NOVICE=.FALSE.
       FAST=.FALSE.
       WRITE(6,102)' Do you want guidance? (Y/N) '
 102   FORMAT(A)
       READ(5,103)CANS
 103   FORMAT(A)
       IF (CANS.EQ.'Y'.OR.CANS.EQ.'y'.OR.CANS.EQ.' ') NOVICE=.TRUE.
       IF (CANS.EQ.'F'.OR.CANS.EQ.'f')  FAST=.TRUE.
```

```
      WRITE(6,11)' What is the name of the file? '
11    FORMAT(A)
      IF (NOVICE) THEN
      WRITE(6,601)' (File containing 3D data)'
601   FORMAT(A)
      ENDIF
      READ(5,12)INFILE
12    FORMAT(A15)
      WRITE(6,13)' What is the scale factor? '
13    FORMAT(A)
      IF (NOVICE) THEN
      WRITE(6,602)' (Size of polygon representing surface at each point'
602   FORMAT(A)
      WRITE(6,603)' Try 1.0 first)'
603   FORMAT(A)
      ENDIF
      READ(*,*)DIAG
      WRITE(6,14)' What orientation do you want to use? '
14    FORMAT(A)
      IF (NOVICE) THEN
      WRITE(6,604)' (X,Y,Z, rotations in degrees)'
604   FORMAT(A)
      ENDIF
      READ(*,*)ROTX,ROTY,ROTZ
      WRITE(6,15)' What segment do you want to use? '
15    FORMAT(A)
      IF (NOVICE) THEN
      WRITE(6,605)' (Current segment number)'
605   FORMAT(A)
      ENDIF
      READ(*,*)ISEG
      WRITE(6,16)' What values of D and S do you want?'
16    FORMAT(A)
      IF (NOVICE) THEN
      WRITE(6,606)' (D = distance from eye to screen'
606   FORMAT(A)
      WRITE(6,607)' A large distance gives a near-orthogonal projection'
607   FORMAT(A)
      WRITE(6,608)' A reasonable value is usually between 100.
     + and 1000.'
608   FORMAT(A)
      WRITE(6,609)' S = screen quotient'
609   FORMAT(A)
      WRITE(6,610)' Scales picture, smaller value gives larger picture'
610   FORMAT(A)
      WRITE(6,611)' Try a value of 1.0 first)'
611   FORMAT(A)
      ENDIF
      READ(*,*)D,S
      WRITE(6,17)' Where do you want to look from?'
17    FORMAT(A)
      IF (NOVICE) THEN
      WRITE(6,612)' (X,Y coordinates of viewpoint: Z is always 0.0'
612   FORMAT(A)
      WRITE(6,613)' This is usually 50.0,50.0; 0.0-100.0 acceptable)'
613   FORMAT(A)
      ENDIF
      READ(*,*)XE,YE
      WRITE(6,18)' What is the light vector?'
```

```
18    FORMAT(A)
      IF (NOVICE) THEN
      WRITE(6,614)' (Direction vector of light rays:'
614   FORMAT(A)
      WRITE(6,615)' It is a left-handed coordinate system)'
615   FORMAT(A)
      ENDIF
      READ(*,*)XL,YL,ZL
      WRITE(6,19)' Do you want squares or octagons? (S/O)'
19    FORMAT(A)
      IF (NOVICE) THEN
      WRITE(6,616)' (Shape of polygon representing surface at
     + each point'
616   FORMAT(A)
      WRITE(6,617)' Type S or O)'
617   FORMAT(A)
      ENDIF
      READ(5,20)CANS
20    FORMAT(A)
      IF (CANS.EQ.'O'.OR.CANS.EQ.'o') THEN
      OCTAGON = .TRUE.
      ELSE
      OCTAGON = .FALSE.
      ENDIF
      WRITE(6,21)' Do you want to draw backward-facing shapes?'
21    FORMAT(A)
      IF (NOVICE) THEN
      WRITE(6,618)' (This is useful to obscure atoms'
618   FORMAT(A)
      WRITE(6,619)' BUT it takes twice as long'
619   FORMAT(A)
      WRITE(6,620)' Type Y or N)'
620   FORMAT(A)
      ENDIF
      READ(5,22)CANS
22    FORMAT(A)
      IF (CANS.EQ.'Y'.OR.CANS.EQ.'y') THEN
      BACKING = .TRUE.
      ELSE
      BACKING = .FALSE.
      ENDIF
      WRITE(6,23)' Please wait...'
23    FORMAT(//,A,/)
CE    Set up contour colours.
      CALL SETCOL
      IF (FAST) GO TO 1
CE    Select view 1 (for picture).
      IVIEW=1
      CALL INT2CH(IVIEW,CIVIEW)
      CALL REAL2CH(DIAG,CDIAG)
      CALL REAL2CH(ROTX,CROTX)
      CALL REAL2CH(ROTY,CROTY)
      CALL REAL2CH(ROTZ,CROTZ)
      CORIENT=CROTX//CROTY//CROTZ
      CALL INT2CH(ISEG,CISEG)
      CALL REAL2CH(D,CD)
      CALL REAL2CH(S,CS)
      CALL REAL2CH(XE,CXE)
      CALL REAL2CH(YE,CYE)
```

```
      CVIEW=CXE//CYE
      CALL REAL2CH(XL,CXL)
      CALL REAL2CH(YL,CYL)
      CALL REAL2CH(ZL,CZL)
      CLIGHT=CXL//CYL//CZL
      CALL CEBORD(1)
      CALL CEDALN(5)
 1    TIMER=.TRUE.
      NCOUNT=0
      CALL CESLVW(IVIEW)
CE Set up menu
      CALL MENUB(IVIEW)
      CALL MENUFILL
CE  Draw "timer" to show progress through program
      CALL CEOPSG(100)
      CALL CEMOVE(100,100)
      CALL CELNIN(1)
      CALL CEDRAW(200,100)
      CALL CEDRAW(200,200)
      CALL CEDRAW(100,200)
      CALL CEDRAW(100,100)
      CALL CELNIN(5)
      CALL CEMOVE(150,150)
      CALL CEDRAW(150,200)
      OPEN(2,FILE=INFILE,STATUS='OLD',
     + CARRIAGECONTROL='NONE')
CE  Remember filename in CFILE.
      CFILE=INFILE
 10   NCOUNT=NCOUNT+1
      READ(2,*)ICOL(NCOUNT),COORD(1,NCOUNT),
     + COORD(2,NCOUNT),COORD(3,NCOUNT),
     + PERP(1,NCOUNT),PERP(2,NCOUNT),PERP(3,NCOUNT)
      IF (ICOL(NCOUNT).LE.0) GO TO 2
      ICOL(NCOUNT)=-ICOL(NCOUNT)
      GO TO 10
 2    CLOSE(2)
CE Change "timer"
      CALL CELNIN(0)
      CALL CEDRAW(150,150)
      CALL CELNIN(5)
      CALL CEDRAW(200,150)
      NCOUNT=NCOUNT-1
      CALL INT2CH(NCOUNT,CNCOUNT)
CE  Transform
 3    CONTINUE
      CALL SETROT(SCALE,ROTX,ROTY,ROTZ)
CE Change "timer"
      IF(ROTATION) CALL CEBNSG(100)
      CALL CELNIN(1)
      CALL CEMOVE(100,100)
      CALL CEDRAW(200,100)
      CALL CEDRAW(200,200)
      CALL CEDRAW(100,200)
      CALL CEDRAW(100,100)
      CALL CEMOVE(200,150)
      CALL CELNIN(0)
      CALL CEDRAW(150,150)
      CALL CELNIN(5)
      CALL CEDRAW(150,100)
```

```
      DO 4 K=1,NCOUNT
      ICOLT(K)=ICOL(K)
      ORIGIN=.TRUE.
      CALL TRANSF(COORD(1,K),COORD(2,K),COORD(3,K),
     + TRANS(1,K),TRANS(2,K),TRANS(3,K),ORIGIN)
      ORIGIN=.FALSE.
      CALL TRANSF(PERP(1,K),PERP(2,K),PERP(3,K),
     + PERPT(1,K),PERPT(2,K),PERPT(3,K),ORIGIN)
 4    CONTINUE
CE If we have drawn a "timer", change it, otherwise carry on
      IF (TIMER) THEN
      CALL CELNIN(0)
      CALL CEDRAW(150,150)
      CALL CELNIN(5)
      CALL CEDRAW(100,150)
      ENDIF
      CALL CELNIN(1)
CE  Order coordinates
      TRANS(3,NCOUNT+1)=999999.0
CE  ORDER3 needs a large number at the end of the array to be sorted
      NT=NCOUNT
      IONE=1
      CALL ORDER3(TRANS,IONE,NT,PERPT,ICOLT)
CE Get rid of "timer"
      CALL CECLSG
      CALL CEDLSG(100)
      TIMER=.FALSE.
CE  Draw picture
 41   CONTINUE
      CALL CEPVSG(IPX,IPY)
      CALL CEDLSG(ISEG)
      CALL CEOPSG(ISEG)
      IF (OCTAGON) THEN
      DO 5 K1=1,NCOUNT
CE  This ensures that only the forward facing octagons are drawn:
      IF (PERPT(3,K1).LT.0) THEN
      IF (.NOT.BACKING) GO TO 5
CE  Draw octagon with colour index 14
      CALL OCTS(TRANS(1,K1),TRANS(2,K1),TRANS(3,K1),PERPT(1,K1),
     + PERPT(2,K1),PERPT(3,K1),XE,YE,0.,0.,-1.,I14,D,S,DIAG)
      ELSE
CE  Draw octagon with colour index ICOLT(K1)
      CALL OCTS(TRANS(1,K1),TRANS(2,K1),TRANS(3,K1),PERPT(1,K1),
     + PERPT(2,K1),PERPT(3,K1),XE,YE,XL,YL,ZL,ICOLT(K1),D,S,DIAG)
      ENDIF
 5    CONTINUE
      ELSE
      DO 6 K2=1,NCOUNT
CE  This ensures that only the forward-facing squares are drawn,
      IF (PERPT(3,K2).LT.0) THEN
CE  unless the backward-facing ones are supposed to be drawn.
      IF (.NOT.BACKING) GO TO 6
CE  Draw square with colour index 14.
      CALL SQUARES(TRANS(1,K2),TRANS(2,K2),TRANS(3,K2),PERPT(1,K2),
     + PERPT(2,K2),PERPT(3,K2),XE,YE,0.,0.,-1.,I14,D,S,DIAG)
      ELSE
CE  Draw square with colour index ICOLT(K2).
      CALL SQUARES(TRANS(1,K2),TRANS(2,K2),TRANS(3,K2),PERPT(1,K2),
     + PERPT(2,K2),PERPT(3,K2),XE,YE,XL,YL,ZL,ICOLT(K2),D,S,DIAG)
```

```
      ENDIF
 6    CONTINUE
      ENDIF
      CALL CECLSG
      IF (FAST) GO TO 99
CE  Ask whether victim has had enough.
      WRITE(6,71)' Do you want to change any parameters?'
 71   FORMAT(//,A,/)
      READ(5,710)CANS
 710  FORMAT(A)
      IF (CANS.EQ.'N'.OR.CANS.EQ.'n') GOTO 99
      CALL CURSOR
      WRITE(6,81)' Calculating new picture - please wait...'
 81   FORMAT(//,A,/)
      CALL CEDLVW(IVIEW)
      CALL CESLVW(IVIEW)
      CALL CEVWPT(0,500,3200,2900)
      CALL CEWIND(0,0,4095,3071)
      CALL CEVISG(-1,0)
      CALL CEPAGE
CE  Check to see if filename has been changed.
      IF (CFILE.NE.INFILE) GO TO 1
CE  Check to see if rotation matrix is to be changed.
      IF (ROTATION) GO TO 3
CE  Otherwise, just re-draw picture.
      GO TO 41
 99   CONTINUE
      CALL CEDALN(32)
      CALL CEGRGN(0,0,0)
      CALL CECRGN(0,0)
      END

      BLOCKDATA WINBLOC
      COMMON /WINDAT/XLEFT,XRIGHT,YBOTTOM,YTOP,XSCALE,YSCALE,
     + XDIFF,YDIFF
      DATA XLEFT,XRIGHT,YBOTTOM,YTOP/0.,3000.,0.,3000./
      DATA XSCALE,YSCALE/1.,1./
      DATA XDIFF,YDIFF/3000.,3000./
      END
```

```
      SUBROUTINE ORDER3(A,P,Q,PERPT,ICOLT)
CE
CE            ORDER3
CE
CE           a version of
CE            QUICKSORT2
CE
CE         from Horowitz & Sahni
CE "Fundamentals of Computer Algorithms"
CE         Computer Science Press
CE
CE  This subroutine sorts the elements A(P),...,A(Q) which
CE  are part of the array A(4000) into ascending order.
CE  N.B. A(P+1) MUST be defined, and must be >= all
CE  elements in A(P:Q).
CE
CE  It is modified to swap the corresponding values for
CE  the x- and y-coordinates, the colour indices and
CE  the perpendicular vector, based on ascending z-coordinates.
CE
        INTEGER TOP,J,P,Q
        INTEGER STACK(30),ICOLT(4000)
        REAL A(3,4000),PERPT(3,4000)
        J=0
        TOP=0
  100   CONTINUE
  200   IF (P.GE.Q) GO TO 300
        J=Q+1
        CALL PARTITION(A,P,J,PERPT,ICOLT)
        IF ((J-P).LT.(Q-J)) THEN
        STACK(TOP+1)=J+1
        STACK(TOP+2)=Q
        Q=J-1
        ELSE
        STACK(TOP+1)=P
        STACK(TOP+2)=J-1
        P=J+1
        ENDIF
        TOP=TOP+2
        GO TO 200
  300   IF (TOP.EQ.0) RETURN
        Q=STACK(TOP)
        P=STACK(TOP-1)
        TOP=TOP-2
        GO TO 100
        END
CE
```

```
      SUBROUTINE PARTITION(A,M,N,PERPT,ICOLT)
CE
CE          PARTITION
CE
CE    procedure from Horowitz & Sahni (see QSORT2)
CE   "Within A(m),A(m+1),...,A(n-1) the elements are
CE    rearranged in such a way that if initially
CE    t=A(m), then after completion A(q)=t, for some q
CE    between m and n-1, A(k).LE.t for m.LE.k.LT.q and
CE    A(k).GE.t for q.LT.k.LT.n
CE    The final value of n is q."
CE
      INTEGER M,N,I,I1
      INTEGER ICOLT(4000)
      REAL A(3,4000),PERPT(3,4000)
      REAL V,V1,V2,P1,P2,P3
      V=A(3,M)
      V1=A(1,M)
      V2=A(2,M)
      P1=PERPT(1,M)
      P2=PERPT(2,M)
      P3=PERPT(3,M)
      I1=ICOLT(M)
      I=M
  100 CONTINUE
  200 I=I+1
      IF (A(3,I).LT.V) GO TO 200
  300 N=N-1
      IF (A(3,N).GT.V) GO TO 300
      IF (I.LT.N) THEN
      CALL INTERCHANGE(A,I,N,PERPT,ICOLT)
      ELSE
      GO TO 400
      ENDIF
      GO TO 100
  400 A(3,M)=A(3,N)
      A(1,M)=A(1,N)
      A(2,M)=A(2,N)
      PERPT(1,M)=PERPT(1,N)
      PERPT(2,M)=PERPT(2,N)
      PERPT(3,M)=PERPT(3,N)
      ICOLT(M)=ICOLT(N)
      A(3,N)=V
      A(1,N)=V1
      A(2,N)=V2
      PERPT(1,N)=P1
      PERPT(2,N)=P2
      PERPT(3,N)=P3
      ICOLT(N)=I1
      RETURN
      END
```

```
      SUBROUTINE INTERCHANGE(A,I,N,PERPT,ICOLT)
CE
CE     INTERCHANGE
CE
      REAL A(3,4000),PERPT(3,4000)
      REAL A1TEMP,A2TEMP,A3TEMP,P1TEMP,P2TEMP,P3TEMP
      INTEGER ICOLT(4000),I,N
      INTEGER ICTEMP
      A1TEMP=A(1,I)
      A2TEMP=A(2,I)
      A3TEMP=A(3,I)
      P1TEMP=PERPT(1,I)
      P2TEMP=PERPT(2,I)
      P3TEMP=PERPT(3,I)
      ICTEMP=ICOLT(I)
      A(1,I)=A(1,N)
      A(2,I)=A(2,N)
      A(3,I)=A(3,N)
      PERPT(1,I)=PERPT(1,N)
      PERPT(2,I)=PERPT(2,N)
      PERPT(3,I)=PERPT(3,N)
      ICOLT(I)=ICOLT(N)
      A(1,N)=A1TEMP
      A(2,N)=A2TEMP
      A(3,N)=A3TEMP
      PERPT(1,N)=P1TEMP
      PERPT(2,N)=P2TEMP
      PERPT(3,N)=P3TEMP
      ICOLT(N)=ICTEMP
      RETURN
      END

      SUBROUTINE WINDOW(XL,YB,XR,YT)
      COMMON /WINDAT/XLEFT,XRIGHT,YBOTTOM,YTOP,XSCALE,YSCALE,
     + XDIFF,YDIFF
      XLEFT=XL
      XRIGHT=XR
      YBOTTOM=YB
      YTOP=YT
      XSCALE=XDIFF/(XRIGHT-XLEFT)
      YSCALE=YDIFF/(YTOP-YBOTTOM)
      RETURN
      END

      SUBROUTINE TRANSF(X,Y,Z,XG,YG,ZG,ORIGIN)
      COMMON /TRANSFORM/SCL(4,4),RX(4,4),RY(4,4),RZ(4,4),
     + UNIT(4,4),T(4,4),W(4,4)
      DIMENSION V(4,1),S(4,1)
      LOGICAL ORIGIN
      CALL AZERO(V,4,1)
      IF (ORIGIN) THEN
CE Move origin to (0,0,0) from (50,50,50)
      V(1,1)=X-50.
      V(2,1)=Y-50.
      V(3,1)=Z-50.
      V(4,1)=1.
      CALL MXMPY(W,V,S,4,4,1)
```

```
      XG=S(1,1)+50.
      YG=S(2,1)+50.
      ZG=S(3,1)+50.
      RETURN
      ELSE
CE Transform without moving origin
      V(1,1)=X
      V(2,1)=Y
      V(3,1)=Z
      V(4,1)=1.
      CALL MXMPY(W,V,S,4,4,1)
      XG=S(1,1)
      YG=S(2,1)
      ZG=S(3,1)
      RETURN
      ENDIF
      END

      SUBROUTINE SETROT(SCALE,ROTX,ROTY,ROTZ)
      COMMON/TRANSFORM/SCL(4,4),RX(4,4),RY(4,4),RZ(4,4),
     + UNIT(4,4),T(4,4),W(4,4)
      REAL PI
      PI=3.1415926
      ROTX=ROTX*PI/180.
      ROTY=ROTY*PI/180.
      ROTZ=ROTZ*PI/180.
      CALL AZERO(UNIT,4,4)
      UNIT(1,1)=1.
      UNIT(2,2)=1.
      UNIT(3,3)=1.
      UNIT(4,4)=1.
      CALL ACOPY(UNIT,RX,4,4)
      RX(2,2)=COS(ROTX)
      RX(3,3)=RX(2,2)
      RX(3,2)=SIN(ROTX)
      RX(2,3)=-RX(3,2)
      CALL ACOPY(UNIT,RY,4,4)
      RY(1,1)=COS(ROTY)
      RY(3,3)=RY(1,1)
      RY(3,1)=SIN(ROTY)
      RY(1,3)=-RY(3,1)
      CALL ACOPY(UNIT,RZ,4,4)
      RZ(1,1)=COS(ROTZ)
      RZ(2,2)=RZ(1,1)
      RZ(2,1)=SIN(ROTZ)
      RZ(1,2)=-RZ(2,1)
      CALL AZERO(SCL,4,4)
      SCL(1,1)=SCALE
      SCL(2,2)=SCALE
      SCL(3,3)=SCALE
      SCL(4,4)=1.
      CALL MXMPY(RY,RX,W,4,4,4)
      CALL MXMPY(RZ,W,T,4,4,4)
      CALL MXMPY(SCL,T,W,4,4,4)
      RETURN
      END
```

```
      SUBROUTINE AZERO(M,IR,IC)
      REAL M(IR,IC)
      DO 20 I=1,IR
      DO 10 J=1,IC
      M(I,J)=0.
10    CONTINUE
20    CONTINUE
      RETURN
      END

      SUBROUTINE ACOPY(S,D,IR,IC)
      DIMENSION S(IR,IC),D(IR,IC)
      DO 20 I=1,IR
      DO 10 J=1,IC
      D(I,J)=S(I,J)
10    CONTINUE
20    CONTINUE
      RETURN
      END

      SUBROUTINE MXMPY(X,Y,Z,II,JJ,KK)
      DIMENSION X(II,JJ),Y(JJ,KK),Z(II,KK)
      DO 30 I=1,II
      DO 20 K=1,KK
      Z(I,K)=0.0
      DO 10 J=1,JJ
      Z(I,K)=Z(I,K)+X(I,J)*Y(J,K)
10    CONTINUE
20    CONTINUE
30    CONTINUE
      RETURN
      END


      SUBROUTINE SQUARES(X,Y,Z,XP,YP,ZP,XE,YE,XL,YL,ZL,INDEX,D,S,DIAG)
CE    X,Y,Z COORDINATES OF POINT
CE   XP,YP,ZP UNIT VECTOR PERPENDICULAR TO SURFACE
CE    AT THAT POINT
CE   XE,YE COORDINATES OF VIEWPOINT
CE   ARE (XE,YE,0)
CE   D IS DISTANCE FROM VIEWPOINT TO SCREEN
CE   S IS SCREEN SIZE
CE   XL,YL,ZL LIGHT VECTOR
CE
CE    THIS ROUTINE DRAWS A COLOURED SQUARE AT (X,Y,Z)
CE    ANGLED ACCORDING TO VECTOR (XP,YP,ZP)
CE
      COMMON /WINDAT/ XLEFT,XRIGHT,YBOTTOM,YTOP,XSCALE,YSCALE,
     + XDIFF,YDIFF
      COMMON /PERPDAT/ V,XPYPDV,XPZPDV,YPDV,ZPDV,XINC,YINC
      REAL X,Y
      REAL XP,YP,ZP
      REAL XL,YL,ZL
      REAL XE,YE
      REAL DIAG,DSIZE
      REAL D,S,DOTPL
      INTEGER*2 INDEX
      INTEGER*2 IX,IY,IVXL,IVYB
      REAL V,XPYPDV,ZPDV,XPZPDV,YPDV
```

```
      REAL XINC,YINC
      REAL X2,Y2,Z2
      LOGICAL SHADE
      IF (YP.EQ.0.AND.ZP.EQ.0) RETURN
CE ZP should be reversed, to mirror reversal of Z
      ZP=-1.*ZP
      SHADE=.FALSE.
      DSIZE=50.*DIAG
CE      DOTPL=(XP*XL)+(YP*YL)+(ZP*ZL)
CE I would have preferred to use the above line,
CE but the stupid compiler thinks it is too
CE complicated!
      DOTPL=(XP*XL)
      DOTPL=DOTPL+(YP*YL)
      DOTPL=DOTPL+(ZP*ZL)
      IF (DOTPL.GT.0.0) SHADE=.TRUE.
      DS = D/S
      IF (SHADE) THEN
      CALL CESFIL(-15)
      GO TO 10
      ENDIF
      CALL CESFIL(INDEX)
 10   CONTINUE
      XINC = -DSIZE/XSCALE
      YINC = -DSIZE/YSCALE
      V = SQRT(YP*YP + ZP*ZP)
      XPYPDV=(XP*YP)/V
      XPZPDV=(XP*ZP)/V
      YPDV=YP/V
      ZPDV=ZP/V
      CALL PERP(X,Y,Z,X2,Y2,Z2,XE,YE,D,DS)
      IX = INT(X2)
      IY = INT(Y2)
      IBORD=0
      CALL CEMOVE(IX,IY)
      CALL CEBPNL(IX,IY,IBORD)
      XINC = DSIZE/XSCALE
      CALL PERP(X,Y,Z,X2,Y2,Z2,XE,YE,D,DS)
      IX = INT(X2)
      IY = INT(Y2)
      CALL CEDRAW(IX,IY)
      YINC = DSIZE/YSCALE
      CALL PERP(X,Y,Z,X2,Y2,Z2,XE,YE,D,DS)
      IX = INT(X2)
      IY = INT(Y2)
      CALL CEDRAW(IX,IY)
      XINC = -DSIZE/XSCALE
      CALL PERP(X,Y,Z,X2,Y2,Z2,XE,YE,D,DS)
      IX = INT(X2)
      IY = INT(Y2)
      CALL CEDRAW(IX,IY)
      CALL CEEPNL
CE Change ZP back.
      ZP=-1.*ZP
      RETURN
      END
```

```
      SUBROUTINE OCTS(X,Y,Z,XP,YP,ZP,XE,YE,XL,YL,ZL,INDEX,D,S,DIAG)
CE    X,Y,Z COORDINATES OF POINT
CE   XP,YP,ZP UNIT VECTOR PERPENDICULAR TO SURFACE
CE    AT THAT POINT
CE   XE,YE COORDINATES OF VIEWPOINT
CE   ARE (XE,YE,0)
CE   D IS DISTANCE FROM VIEWPOINT TO SCREEN
CE   S IS SCREEN SIZE
CE   XL,YL,ZL LIGHT VECTOR
CE
CE     THIS ROUTINE DRAWS A COLOURED OCTAGON AT (X,Y,Z)
CE    ANGLED ACCORDING TO VECTOR (XP,YP,ZP)
CE
      COMMON /WINDAT/ XLEFT,XRIGHT,YBOTTOM,YTOP,XSCALE,YSCALE,
     + XDIFF,YDIFF
      COMMON /PERPDAT/ V,XPYPDV,XPZPDV,YPDV,ZPDV,XINC,YINC
      REAL X,Y
      REAL XP,YP,ZP
      REAL XL,YL,ZL
      REAL XE,YE
      REAL D,S,DOTPL,PID8,RAD,DIAG
      REAL SPID8X,CPID8X
      REAL SPID8Y,CPID8Y
      INTEGER*2 INDEX
      INTEGER*2 IX,IY,IVXL,IVYB
      REAL V,XPYPDV,ZPDV,XPZPDV,YPDV
      REAL XINC,YINC
      REAL X2,Y2,Z2
      LOGICAL SHADE
      IF (YP.EQ.0.AND.ZP.EQ.0) RETURN
CE ZP should be reversed, to mirror Z reversal
      ZP=-1.*ZP
      SHADE=.FALSE.
      PID8=3.14159/8.
      RAD=50.*DIAG
      SPID8X=RAD*SIN(PID8)/XSCALE
      SPID8Y=RAD*SIN(PID8)/YSCALE
      CPID8X=RAD*COS(PID8)/XSCALE
      CPID8Y=RAD*COS(PID8)/YSCALE
CE      DOTPL=(XP*XL)+(YP*YL)+(ZP*ZL)
CE I would have preferred to use the above line,
CE but the stupid compiler thinks it is too
CE complicated!
      DOTPL=(XP*XL)
      DOTPL=DOTPL+(YP*YL)
      DOTPL=DOTPL+(ZP*ZL)
      IF (DOTPL.GT.0.0) SHADE=.TRUE.
      DS = D/S
      IF (SHADE) THEN
      CALL CESFIL(-15)
      GO TO 10
      ENDIF
      CALL CESFIL(INDEX)
 10   CONTINUE
      XINC = -SPID8X
      YINC = -CPID8Y
      V = SQRT(YP*YP + ZP*ZP)
      XPYPDV=(XP*YP)/V
      XPZPDV=(XP*ZP)/V
```

```
      YPDV=YP/V
      ZPDV=ZP/V
      CALL PERP(X,Y,Z,X2,Y2,Z2,XE,YE,D,DS)
      IX = INT(X2)
      IY = INT(Y2)
      IBORD=0
      CALL CEMOVE(IX,IY)
      CALL CEBPNL(IX,IY,IBORD)
      XINC =  SPID8X
      YINC = -CPID8Y
      CALL PERP(X,Y,Z,X2,Y2,Z2,XE,YE,D,DS)
      IX = INT(X2)
      IY = INT(Y2)
      CALL CEDRAW(IX,IY)
      XINC =  CPID8X
      YINC = -SPID8Y
      CALL PERP(X,Y,Z,X2,Y2,Z2,XE,YE,D,DS)
      IX = INT(X2)
      IY = INT(Y2)
      CALL CEDRAW(IX,IY)
      XINC = CPID8X
      YINC = SPID8Y
      CALL PERP(X,Y,Z,X2,Y2,Z2,XE,YE,D,DS)
      IX = INT(X2)
      IY = INT(Y2)
      CALL CEDRAW(IX,IY)
      XINC = SPID8X
      YINC = CPID8Y
      CALL PERP(X,Y,Z,X2,Y2,Z2,XE,YE,D,DS)
      IX = INT(X2)
      IY = INT(Y2)
      CALL CEDRAW(IX,IY)
      XINC = -SPID8X
      YINC =  CPID8Y
      CALL PERP(X,Y,Z,X2,Y2,Z2,XE,YE,D,DS)
      IX = INT(X2)
      IY = INT(Y2)
      CALL CEDRAW(IX,IY)
      XINC = -CPID8X
      YINC =  SPID8Y
      CALL PERP(X,Y,Z,X2,Y2,Z2,XE,YE,D,DS)
      IX = INT(X2)
      IY = INT(Y2)
      CALL CEDRAW(IX,IY)
      XINC = -CPID8X
      YINC = -SPID8Y
      CALL PERP(X,Y,Z,X2,Y2,Z2,XE,YE,D,DS)
      IX = INT(X2)
      IY = INT(Y2)
      CALL CEDRAW(IX,IY)
      CALL CEEPNL
CE Change ZP back again.
      ZP=-1.*ZP
      RETURN
      END
```

```
      SUBROUTINE PERP(X,Y,Z,X2,Y2,Z2,XE,YE,D,DS)
CE Input:
CE       X,Y,Z -  centre of coloured area
CE       D - distance to observer
CE       S - screen scale quotient
CE       XINC,YINC - increments from X,Y to new point
CE       position of observer is (XE,YE,0)
CE
CE Output:
CE       X2,Y2,Z2 - points after perspective transformation.
CE
      COMMON /WINDAT/ XLEFT,XRIGHT,YBOTTOM,YTOP,XSCALE,YSCALE,
     + XDIFF,YDIFF
      COMMON /PERPDAT/ V,XPYPDV,XPZPDV,YPDV,ZPDV,XINC,YINC
      REAL X,Y,Z
      REAL XE,YE
      REAL D,S
      REAL V,XPYPDV,XPZPDV,YPDV,ZPDV
      REAL XINC,YINC
      REAL X2,Y2,Z2
      INTEGER*2 IVXL,IVYB
      IVXL=0
      IVYB=0
      Z2 = -XPZPDV*XINC - YPDV*YINC + Z
CE Reverse Z-coordinate direction
      Z2 = -1.*Z2
      Z2 = Z2 + D
      X2 = V*XINC + X - XE
      X2 = DS*X2/Z2
      X2 = X2 + XE
      X2 = (X2-XLEFT)*XSCALE + IVXL
      Y2 = -XPYPDV*XINC + ZPDV*YINC + Y - YE
      Y2 = DS*Y2/Z2
      Y2 = Y2 + YE
      Y2 = (Y2-YBOTTOM)*YSCALE + IVYB
      RETURN
      END

      SUBROUTINE REAL2CH(REIN,CHARVAL)
CE    This routine converts real numbers to character*6
      REAL REIN
      CHARACTER*6 CHARVAL
      OPEN (UNIT=1,STATUS='SCRATCH',ACCESS='SEQUENTIAL',
     + FORM='FORMATTED')
      WRITE(1,501)REIN
 501  FORMAT(F6.1)
      REWIND(1)
      READ(1,502)CHARVAL
 502  FORMAT(A6)
      CLOSE(1)
      RETURN
      END

      SUBROUTINE INT2CH(NOIN,CHARVAL)
CE    This routine converts integers into character*5
      INTEGER NOIN
      CHARACTER*5 CHARVAL
      OPEN (UNIT=1,STATUS='SCRATCH',ACCESS='SEQUENTIAL',
     + FORM='FORMATTED')
      WRITE(1,501)NOIN
 501  FORMAT(I5)
```

```
      REWIND(1)
      READ(1,502)CHARVAL
502   FORMAT(A5)
      CLOSE(1)
      RETURN
      END

      SUBROUTINE MENUB(IVIEW)
CE    This subroutine draws the border for the menu,
CE    as well as all the headings.
CE    Select view IVIEW for the picture.
        CALL CESLVW(IVIEW)
CE    Viewport for picture.
        CALL CEVWPT(0,500,3200,2900)
CE    Window for picture.
        CALL CEWIND(0,0,4095,3071)
CE    Segment pivot point.
        CALL CEPVSG(3200,0)
CE    Select view 64 for the menu area.
        CALL CESLVW(64)
CE    Viewport for menu.
        CALL CEVWPT(3200,0,4095,2900)
CE    Window to address this viewport.
        CALL CEWIND(3200,0,4095,2900)
CE    Open a segment to store the menu in.
        CALL CEDLSG(500)
        CALL CEBNSG(500)
CE    Set graphtext size and colour.
CE    These values are default text size and colour yellow.
        CALL CEGSIZ(39,59,12)
        CALL CETXIN(7)
CE    Write header.
        CALL CEMOVE(3500,2800)
        CALL CETEXT(4,'MENU')
CE    Set graphtext size and colour for variables' headings.
        CALL CEGSIZ(30,45,10)
        CALL CETXIN(2)
CE    Write out headings.
        CALL CEMOVE(3250,2700)
        CALL CETEXT(8,'Filename')
        CALL CEMOVE(3250,2500)
        CALL CETEXT(11,'Orientation')
        CALL CEMOVE(3250,2300)
        CALL CETEXT(13,'View Distance ')
        CALL CEMOVE(3250,2100)
        CALL CETEXT(15,'Screen Quotient')
        CALL CEMOVE(3250,1900)
        CALL CETEXT(5,'Light')
        CALL CEMOVE(3250,1700)
        CALL CETEXT(9,'Viewpoint')
        CALL CEMOVE(3250,1500)
        CALL CETEXT(13,'Polygon Shape')
        CALL CEMOVE(3250,1300)
        CALL CETEXT(12,'Polygon Size')
        CALL CEMOVE(3250,1100)
        CALL CETEXT(7,'Backing')
        CALL CEMOVE(3250,900)
        CALL CETEXT(14,'Segment Number')
        CALL CEMOVE(3250,700)
```

```
      CALL CETEXT(11,'View number')
      CALL CETXIN(7)
      CALL CEGSIZ(39,59,12)
      CALL CEMOVE(3500,430)
      CALL CETEXT(4,'EXIT')
      CALL CEMOVE(3500,230)
      CALL CETEXT(4,'HELP')
      CALL CECLSG
CE  Create segment for the cursor.
      CALL CELNIN(3)
      CALL CEDLSG(501)
      CALL CEBNSG(501)
      CALL CEMOVE(0,200)
      CALL CEDRAW(4095,200)
      CALL CEMOVE(4095,400)
      CALL CEDRAW(0,400)
      CALL CECLSG
      CALL CEGRGN(0,0,200)
      CALL CECRGN(0,501)
      CALL CELNIN(1)
      CALL CESLVW(IVIEW)
      CALL CEPVSG(0,0)
      RETURN
      END

      SUBROUTINE MENUFILL
CE  This routine fills in the menu values
CE  IVIEW is the current view
      COMMON /MENU/INFILE,CORIENT,CD,CS,CLIGHT,CVIEW,CDIAG,
     + CISEG,CIVIEW
      COMMON /VALUES/ROTX,ROTY,ROTZ,D,S,XL,YL,ZL,XE,YE,DIAG,
     + ISEG,IVIEW,OCTAGON,BACKING,ROTATION
      CHARACTER*18 CORIENT,CLIGHT,CINFILE
      CHARACTER*12 CVIEW
      CHARACTER*15 INFILE
      CHARACTER*6 CD,CS,CDIAG
      CHARACTER*10 CSHAPE
      CHARACTER*5 CISEG,CIVIEW
      CHARACTER*6 CBACK
      LOGICAL OCTAGON,BACKING,ROTATION
CE  Select view 64
      CALL CESLVW(64)
CE  Get rid of old values
      CALL CEDLSG(502)
      CALL CEBNSG(502)
CE  Adjust text size
      CALL CEGSIZ(35,45,10)
      CALL CEMOVE(3250,2600)
      CINFILE='   '//INFILE
      CALL CETEXT(18,CINFILE)
      CALL CEMOVE(3250,2400)
      CALL CETEXT(18,CORIENT)
      CALL CEMOVE(3250,2200)
      CALL CETEXT(6,CD)
      CALL CEMOVE(3250,2000)
      CALL CETEXT(6,CS)
      CALL CEMOVE(3250,1800)
      CALL CETEXT(18,CLIGHT)
      CALL CEMOVE(3250,1600)
      CALL CETEXT(12,CVIEW)
```

```
      CALL CEMOVE(3250,1400)
      IF (OCTAGON) THEN
      CSHAPE='   Octagon'
      ELSE
      CSHAPE='   Square '
      ENDIF
      CALL CETEXT(10,CSHAPE)
      CALL CEMOVE(3250,1200)
      CALL CETEXT(6,CDIAG)
      CALL CEMOVE(3250,1000)
      IF (BACKING) THEN
      CBACK= '   Yes'
      ELSE
      CBACK= '   No '
      ENDIF
      CALL CETEXT(6,CBACK)
      CALL CEMOVE(3250,800)
      CALL CETEXT(5,CISEG)
      CALL CEMOVE(3250,600)
      CALL CETEXT(5,CIVIEW)
      CALL CECLSG
      CALL CESLVW(IVIEW)
      CONTINUE
      RETURN
      END

      SUBROUTINE SETCOL
CE This program sets up the colours for the 3d program.
       DIMENSION ICLRY(36)
       DATA ICLRY/1,120,50,100, 2,145,50,100, 3,160,50,100,
     +            4,180,50,100, 5,200,50,100, 6,280,50,100,
     +            7,300,50,100, 8,  0,50,100, 9, 60,50,100/
       CALL CECLMP(1,36,ICLRY)
       END

      SUBROUTINE CURSOR
      COMMON /MENU/INFILE,CORIENT,CD,CS,CLIGHT,CVIEW,CDIAG,
     + CISEG,CIVIEW
      COMMON /VALUES/ROTX,ROTY,ROTZ,D,S,XL,YL,ZL,XE,YE,DIAG,
     + ISEG,IVIEW,OCTAGON,BACKING,ROTATION
      CHARACTER*18 CORIENT,CLIGHT
      CHARACTER*15 INFILE
      CHARACTER*12 CVIEW
      CHARACTER*6 CROTX,CROTY,CROTZ
      CHARACTER*6 CXL,CYL,CZL,CXE,CYE
      CHARACTER*6 CD,CS,CDIAG
      CHARACTER*5 CISEG,CIVIEW
      CHARACTER*1 SIGCHR,KEYCHR,CANS,SIGCHRD,KEYCHRD
      LOGICAL HELP,OCTAGON,BACKING,ROTATION
      HELP = .FALSE.
CE  This saves time by establishing whether there is
CE  a change in the rotation matrix:
      ROTATION = .FALSE.
 100  CONTINUE
      CALL MENUFILL
      CALL CESLVW(64)
      WRITE(6,72)'            CURSOR ACTIVATED'
 72   FORMAT(///,A,/)
      CALL CEMOVE(3500,230)
      CALL CEGSIZ(39,59,12)
```

```
      IF (HELP) THEN
      CALL CETXIN(9)
      ELSE
      CALL CETXIN(7)
      ENDIF
      CALL CETEXT(4,'HELP')
      CALL CETXIN(7)
      CALL CEGSIZ(30,45,10)
CE    CALL WAIT(0.5)
      CALL CEENGN(0,0)
      CALL CEGTGN(SIGCHR,KEYCHR,IX,IY,IDUM1,IDUM2)
CE    CALL WAIT(0.5)
      CALL CEDSGN(0)
      CALL CEGTGN(SIGCHRD,KEYCHRD,IXD,IYD,IDUM3,IDUM4)
      IF (KEYCHR.EQ.'E') GO TO 999
CE  Filename
      IF (IY.EQ.2600) GO TO 28
CE  Orientation
      IF (IY.EQ.2400) GO TO 26
CE  View Distance
      IF (IY.EQ.2200) GO TO 24
CE  Screen Quotient
      IF (IY.EQ.2000) GO TO 22
CE  Light
      IF (IY.EQ.1800) GO TO 20
CE  Viewpoint
      IF (IY.EQ.1600) GO TO 18
CE  Polygon Shape
      IF (IY.EQ.1400) GO TO 16
CE  Polygon Size
      IF (IY.EQ.1200) GO TO 14
CE  Backing
      IF (IY.EQ.1000) GO TO 12
CE  Segment Number
      IF (IY.EQ.800) GO TO 10
CE  View Number
      IF (IY.EQ.600)  GO TO  8
CE  EXIT
      IF (IY.EQ.400)  GO TO  6
CE  Otherwise, Help needed
CE  If HELP is already true (i.e. HELP picked twice in a row)
CE  Give more explicit instructions.
CE      CALL CEDSGN(0)
      IF (HELP) THEN
      WRITE(6,601)' Move cursor by pressing joydisk'
 601  FORMAT(/,A)
      WRITE(6,602)' To pick an item press space bar'
 602  FORMAT(A)
      HELP = .FALSE.
      ELSE
      HELP = .TRUE.
      WRITE(6,603)' Pick item with cursor'
 603  FORMAT(/,A,/)
      ENDIF
      GO TO 100
CE 28  FILENAME
 28   IF (HELP) THEN
      WRITE(6,604)' File with 3D data in it'
 604  FORMAT(/,A)
      ENDIF
```

```
      WRITE(6,605)' Type in new filename'
 605  FORMAT(/,A)
      READ(5,505)INFILE
 505  FORMAT(A15)
      GO TO 100
CE 26   X, Y, Z ROTATIONS
 26   IF (HELP) THEN
      WRITE(6,606)' X,Y,Z rotations in degrees'
 606  FORMAT(/,A)
      ENDIF
      ROTATION = .TRUE.
      WRITE(6,607)' Type in new orientation: ROTX,ROTY,ROTZ'
 607  FORMAT(/,A)
      READ(*,*)ROTX,ROTY,ROTZ
      CALL REAL2CH(ROTX,CROTX)
      CALL REAL2CH(ROTY,CROTY)
      CALL REAL2CH(ROTZ,CROTZ)
      CORIENT=CROTX//CROTY//CROTZ
      GO TO 100
CE 24   D
 24   IF (HELP) THEN
      WRITE(6,608)' Distance from eye to screen'
 608  FORMAT(/,A)
      WRITE(6,609)' A large distance gives a near-orthogonal projection'
 609  FORMAT(A)
      WRITE(6,610)' A reasonable value is usually around 100.0'
 610  FORMAT(A)
      ENDIF
      WRITE(6,611)' Type in new view distance'
 611  FORMAT(/,A)
      READ(*,*)D
      CALL REAL2CH(D,CD)
      GO TO 100
CE 22   S
 22   IF (HELP) THEN
      WRITE(6,612)' Scales picture, smaller value gives larger picture'
 612  FORMAT(/,A)
      WRITE(6,613)' A reasonable value is usually around 0.5'
 613  FORMAT(A)
      ENDIF
      WRITE(6,614)' Type in new screen quotient'
 614  FORMAT(/,A)
      READ(*,*)S
      CALL REAL2CH(S,CS)
      GO TO 100
CE 20   LIGHT
 20   IF (HELP) THEN
      WRITE(6,615)' Direction vector of light rays'
 615  FORMAT(/,A)
      WRITE(6,616)' It is a left-handed coordinate system'
 616  FORMAT(A)
      ENDIF
      WRITE(6,617)' Type in new light vector: XL,YL,ZL'
 617  FORMAT(/,A)
      READ(*,*)XL,YL,ZL
      CALL REAL2CH(XL,CXL)
      CALL REAL2CH(YL,CYL)
      CALL REAL2CH(ZL,CZL)
      CLIGHT=CXL//CYL//CZL
      GO TO 100
```

```
CE 18   VIEWPOINT
 18    IF (HELP) THEN
       WRITE(6,618)' X,Y coordinates of viewpoint: Z is always 0.0'
 618   FORMAT(/,A)
       WRITE(6,619)' This is usually 50.0,50.0'
 619   FORMAT(A)
       ENDIF
       WRITE(6,620)' Type in viewpoint: XE,YE'
 620   FORMAT(/,A)
       READ(*,*)XE,YE
       CALL REAL2CH(XE,CXE)
       CALL REAL2CH(YE,CYE)
       CVIEW=CXE//CYE
       GO TO 100
CE 16   POLYGON SHAPE
 16    IF (HELP) THEN
       WRITE(6,621)' Shape of polygon representing surface at each point'
 621   FORMAT(/,A)
       WRITE(6,622)' This toggles between squares and octagons'
 622   FORMAT(A)
       ENDIF
       WRITE(6,623)' Change of polygon shape'
 623   FORMAT(/,A,/)
       IF (OCTAGON) THEN
       OCTAGON = .FALSE.
       ELSE
       OCTAGON = .TRUE.
       ENDIF
       GO TO 100
CE 14   POLYGON SIZE
 14    IF (HELP) THEN
       WRITE(6,624)' Size of polygon representing surface at each point'
 624   FORMAT(/,A)
       WRITE(6,625)' Try 1.0 first'
 625   FORMAT(A)
       ENDIF
       WRITE(6,626)' Type in new polygon size'
 626   FORMAT(/,A)
       READ(*,*)DIAG
       CALL REAL2CH(DIAG,CDIAG)
       GO TO 100
CE 12   BACKING
 12    IF (HELP) THEN
       WRITE(6,627)' Backward-facing polygons can be displayed'
 627   FORMAT(/,A)
       WRITE(6,628)' This is useful to obscure atoms'
 628   FORMAT(A)
       WRITE(6,629)' BUT it takes twice as long'
 629   FORMAT(A)
       ENDIF
       WRITE(6,630)' Toggle for rear-facing polygons'
 630   FORMAT(/,A,/)
       IF (BACKING) THEN
       BACKING = .FALSE.
       ELSE
       BACKING = .TRUE.
       ENDIF
       GO TO 100
CE 10   SEGMENT NUMBER
 10    IF (HELP) THEN
```

```
      WRITE(6,631)' Current segment number'
  631 FORMAT(/,A)
      ENDIF
      WRITE(6,632)' Type in new segment number'
  632 FORMAT(/,A)
      READ(*,*)ISEG
      CALL CEDLSG(ISEG)
      CALL INT2CH(ISEG,CISEG)
      GO TO 100
CE 8    VIEW NUMBER
  8   IF (HELP) THEN
      WRITE(6,633)' Current view number: 1-63 allowed'
  633 FORMAT(/,A)
      WRITE(6,634)' Try to store segments in different views'
  634 FORMAT(A)
      WRITE(6,635)' It is quicker to change the picture this way'
  635 FORMAT(A)
      ENDIF
      WRITE(6,636)' Type in new view number (NOT 64)'
  636 FORMAT(/,A)
      READ(*,*)IVIEW
      CALL INT2CH(IVIEW,CIVIEW)
      GO TO 100
CE 6    EXIT
  6   IF (HELP) THEN
      WRITE(6,637)' You must be joking! - EXIT is used to exit MENU'
  637 FORMAT(/,A)
      ENDIF
  999 CONTINUE
      CALL CESLVW(IVIEW)
      RETURN
      END

      SUBROUTINE WAIT(SECONDS)
      INTEGER ISTAT
      REAL SECONDS
      ISTAT = LIB$WAIT(SECONDS)
      RETURN
      END
```

Graphics Primitive Subroutines

The following subroutines and functions are the VAX-equivalent versions of the Tektronix DTI routines necessary to run this program:

CEBNSG	BEGIN-NEW-SEGMENT
CEBORD	SET-BORDER-VISIBILITY
CEBPNL	BEGIN-PANEL-BOUNDARY
CECLMP	SET-SURFACE-COLOUR-MAP
CECLSG	CLOSE-SEGMENT
CECRGN	SET-GIN-CURSOR
CEDACL	CLEAR-DIALOGUE-SCROLL
CEDALN	SET-DIALOGUE-AREA-LINES
CEDLSG	DELETE-SEGMENT
CEDLVW	DELETE-VIEW
CEDRAW	DRAW
CEDSGN	DISABLE-GIN
CEENGN	ENABLE-GIN
CEEPNL	END-PANEL
CEGRGN	SET-GIN-GRIDDING
CEGSIZ	SET-GRAPHTEXT-SIZE
CEGTGN	GET-GIN-REPORT
CEINIT	INITIALISE
CELNIN	SET-LINE-INDEX
CEMOVE	MOVE
CEOPSG	BEGIN-SEGMENT
CEPAGE	PAGE
CEPVSG	SET-SEGMENT-PIVOT-POINT
CERPSG	SET-REPORT-SIGNATURE-CHARS
CESFIL	SELECT-FILL-PATTERN
CESLVW	SELECT-VIEW
CETEXT	GRAPHIC-TEXT
CETXIN	SET-TEXT-INDEX
CEVISG	SET-SEGMENT-VISIBILITY
CEVWPT	SET-VIEWPORT
CEWIND	SET-WINDOW
CODEIN	ENCODE-INTEGER-VALUE
CODEXY	ENCODE-XY-VALUES
DECODEXY	DECODE-XY-REPORT

```
      SUBROUTINE CEBNSG(ISEG)
CE Equivalent to LLBNSG
CE BEGIN-NEW-SEGMENT
CE ISEG segment number
      CHARACTER*1 ESC
      CHARACTER*4 CODEIN,INTVAL
      INTEGER ISEG
      ESC=CHAR(27)
      INTVAL=CODEIN(ISEG)
      WRITE(6,60)ESC//'SE'//INTVAL
 60   FORMAT('+',A$)
      END

      SUBROUTINE CEBORD(IVIS)
CE Equivalent to LLBORD
CE SET-BORDER-VISIBILITY
CE IVIS border visibility
CE 0 invisible; 1 visible; 2 toggle
      CHARACTER*1 ESC,RET
      CHARACTER*4 INTSTR,CODEIN
      INTEGER IVIS
      ESC=CHAR(27)
      INTSTR=CODEIN(IVIS)
      WRITE(6,60)ESC//'RE'//INTSTR
 60   FORMAT('+',A$)
      END

      SUBROUTINE CEBPNL(IX,IY,IBOUND)
CE Equivalent to LLBPNL
CE BEGIN-PANEL-BOUNDARY
CE IX,IY coordinates of first point on panel boundary
CE IBOUND 1 draw panel boundary, 0 do not
      CHARACTER*1 ESC,RET
      CHARACTER*5 XYSTR,CODEXY
      CHARACTER*4 INTSTR,CODEIN
      INTEGER IX,IY,IBOUND
      ESC=CHAR(27)
      RET=CHAR(13)
      XYSTR=CODEXY(IX,IY)
      INTSTR=CODEIN(IBOUND)
      WRITE(6,60)ESC//'LP'//XYSTR//INTSTR
 60   FORMAT('+',A$)
      END

      SUBROUTINE CECLMP(ISURF,ILEN,ICLRY)
CE Equivalent to LLCLMP
CE SET-SURFACE-COLOUR-MAP
CE ISURF surface number
CE ILEN length of colour map array ICLRY
CE ICLRY array of colour map data
      INTEGER ISURF,ILEN,J1,J2,J3
      INTEGER ICLRY(ILEN)
      CHARACTER*1 ESC,RET
      CHARACTER*4 CODEIN,INTSTR
      CHARACTER*4 CODE1,CODE2,CODE3,CODE4
      CHARACTER*16 CODE
      ESC=CHAR(27)
      RET=CHAR(13)
      INTSTR=CODEIN(ISURF)
```

```
      J=2
10    IF(INTSTR(J:J).EQ.RET) GOTO 20
      J=J+1
      GOTO 10
20    DO 100 N=1,ILEN,4
      J1=2
      J2=2
      J3=2
      CODE1=CODEIN(ICLRY(N))
      CODE2=CODEIN(ICLRY(N+1))
      CODE3=CODEIN(ICLRY(N+2))
      CODE4=CODEIN(ICLRY(N+3))
90    IF (CODE1(J1:J1).EQ.RET) GOTO 91
      J1=J1+1
      GOTO 90
91    IF (CODE2(J2:J2).EQ.RET) GOTO 92
      J2=J2+1
      GOTO 91
92    IF (CODE3(J3:J3).EQ.RET) GOTO 93
      J3=J3+1
      GOTO 92
93    CODE='4'//CODE1(1:J1-1)//CODE2(1:J2-1)//
     + CODE3(1:J3-1)//CODE4
      WRITE(6,60)ESC//'TG'//INTSTR(1:J-1)//CODE
60    FORMAT('+',A$)
100   CONTINUE
      RETURN
      END

      SUBROUTINE CECLSG
CE Equivalent to LLCLSG
CE CLOSE-SEGMENT
CE ISEG segment number
      CHARACTER*1 ESC,RET
      ESC=CHAR(27)
      RET=CHAR(13)
      WRITE(6,60)ESC//'SC'//RET
60    FORMAT('+',A$)
      END

      SUBROUTINE CECRGN(IDVFNC,ISEG)
CE Equivalent to LLCRGN
CE SET-GIN-CURSOR
CE IDVFNC device function code
CE ISEG segment number
      CHARACTER*1 ESC,RET
      CHARACTER*4 CODEIN,INTVAL1,INTVAL2
      INTEGER IDVFNC,ISEG
      ESC=CHAR(27)
      RET=CHAR(13)
      INTVAL1=CODEIN(IDVFNC)
CE Remove RET from end of string
      J=2
1     IF (INTVAL1(J:J).EQ.RET) GOTO 2
      J=J+1
      GO TO 1
2     INTVAL2=CODEIN(ISEG)
      WRITE(6,60)ESC//'IC'//INTVAL1(1:J-1)//INTVAL2
60    FORMAT('+',A$)
      END
```

```
      SUBROUTINE CEDACL
CE Equivalent to LLDACL
CE CLEAR-DIALOGUE-SCROLL
      CHARACTER*1 ESC,RET
      ESC=CHAR(27)
      RET=CHAR(13)
      WRITE(6,60)ESC//'LZ'//RET
 60   FORMAT('+',A$)
      END

      SUBROUTINE CEDALN(ILINES)
CE Equivalent to LLDALN
CE SET-DIALOGUE-AREA-LINES
CE ILINES number of dialog lines visible
      CHARACTER*1 ESC
      CHARACTER*4 INTSTR,CODEIN
      INTEGER ILINES
      ESC=CHAR(27)
      INTSTR=CODEIN(ILINES)
      WRITE(6,60)ESC//'LL'//INTSTR
 60   FORMAT('+',A$)
      END

      SUBROUTINE CEDLSG(ISEG)
CE Equivalent to LLDLSG
CE DELETE-SEGMENT
CE ISEG segment number
      CHARACTER*1 ESC
      CHARACTER*4 INTSTR,CODEIN
      INTEGER ISEG
      ESC=CHAR(27)
      INTSTR=CODEIN(ISEG)
      WRITE(6,60)ESC//'SK'//INTSTR
 60   FORMAT('+',A$)
      END

      SUBROUTINE CEDLVW(IVIEW)
CE Equivalent to LLDLVW
CE DELETE-VIEW
CE IVIEW view number
      CHARACTER*1 ESC
      CHARACTER*4 INTSTR,CODEIN
      INTEGER IVIEW
      ESC=CHAR(27)
      INTSTR=CODEIN(IVIEW)
      WRITE(6,60)ESC//'RK'//INTSTR
 60   FORMAT('+',A$)
      RETURN
      END

      SUBROUTINE CEDRAW(IX,IY)
CE Equivalent to LLDRAW
CE DRAW
CE IX,IY coordinates to draw to
      CHARACTER*1 ESC,RET
      CHARACTER*5 XYSTR,CODEXY
      INTEGER IX,IY
      ESC=CHAR(27)
      RET=CHAR(13)
```

```
      XYSTR=CODEXY(IX,IY)
      WRITE(6,60)ESC//'LG'//XYSTR//RET
 60   FORMAT('+',A$)
      END

      SUBROUTINE CEDSGN(IDVFNC)
CE Equivalent to LLDSGN
CE DISABLE-GIN
CE IDVFNC device function code
      CHARACTER*1 ESC
      CHARACTER*4 CODEIN,INTVAL
      INTEGER IDVFNC
      ESC=CHAR(27)
      INTVAL=CODEIN(IDVFNC)
      WRITE(6,60)ESC//'ID'//INTVAL
 60   FORMAT(1X,A)
      END

      SUBROUTINE CEENGN(IDVFNC,INUMB)
CE Equivalent to LLENGN
CE ENABLE-GIN
CE IDVFNC device function code
CE INUMB  number of events
      CHARACTER*1 ESC,RET
      CHARACTER*4 CODEIN,CIDVFNC,CINUMB
      INTEGER IDVFNC,INUMB
      ESC=CHAR(27)
      RET=CHAR(13)
      CIDVFNC=CODEIN(IDVFNC)
      N=1
 1    CONTINUE
      IF (CIDVFNC(N+1:N+1).NE.RET) THEN
      N=N+1
      GOTO 1
      ENDIF
      CINUMB=CODEIN(INUMB)
      WRITE(6,60)ESC//'IE'//CIDVFNC(1:N)//CINUMB
 60   FORMAT('+',A$)
      END

      SUBROUTINE CEEPNL
CE Equivalent to LLEPNL
CE END-PANEL
      CHARACTER*1 ESC,RET
      ESC=CHAR(27)
      RET=CHAR(13)
      WRITE(6,60)ESC//'LE'//RET
 60   FORMAT('+',A$)
      END

      SUBROUTINE CEGRGN(IDVFNC,IGRIDX,IGRIDY)
CE Equivalent to LLGRGN
CE SET-GIN-GRIDDING
CE IDVFNC device function code
CE IGRIDX x grid spacing
CE IGRIDY y grid spacing
      CHARACTER*1 ESC,RET
      CHARACTER*4 CODEIN
      CHARACTER*4 CIDVFNC,CIGRIDX,CIGRIDY
```

```
      INTEGER IDVFNC,IGRIDX,IGRIDY,J1,J2
      ESC=CHAR(27)
      RET=CHAR(13)
      CIDVFNC=CODEIN(IDVFNC)
      CIGRIDX=CODEIN(IGRIDX)
      CIGRIDY=CODEIN(IGRIDY)
CE Remove RET from end of string
      J1=2
      J2=2
 1    IF (CIDVFNC(J1:J1).EQ.RET) GOTO 2
      J1=J1+1
      GO TO 1
 2    IF (CIGRIDX(J2:J2).EQ.RET) GOTO 3
      J2=J2+1
      GOTO 2
 3    CONTINUE
 4    CIGRIDY=CODEIN(IGRIDY)
      WRITE(6,60)ESC//'IG'//CIDVFNC(1:J1-1)//
     + CIGRIDX(1:J2-1)//CIGRIDY
 60   FORMAT('+',A$)
      END

      SUBROUTINE CEGSIZ(IWIDTH,IHGHT,ISPACE)
CE Equivalent to LLGSIZ
CE SET-GRAPHTEXT-SIZE
CE IWIDTH width of character cell
CE IHGHT  height of character cell
CE ISPACE spacing between character cells
      CHARACTER*1 ESC,RET
      CHARACTER*4 CODEIN
      CHARACTER*4 INTVAL1,INTVAL2,INTVAL3
      INTEGER IWIDTH,IHGHT,ISPACE,J1,J2,J3
      ESC=CHAR(27)
      RET=CHAR(13)
      INTVAL1=CODEIN(IWIDTH)
      INTVAL2=CODEIN(IHGHT)
      INTVAL3=CODEIN(ISPACE)
CE Remove RET from end of string
      J1=2
      J2=2
 1    IF (INTVAL1(J1:J1).EQ.RET) GOTO 2
      J1=J1+1
      GOTO 1
 2    IF (INTVAL2(J2:J2).EQ.RET) GOTO 3
      J2=J2+1
      GOTO 2
 3    CONTINUE
 4    INTVAL3=CODEIN(ISPACE)
      WRITE(6,60)ESC//'MC'//INTVAL1(1:J1-1)//
     + INTVAL2(1:J2-1)//INTVAL3
 60   FORMAT('+',A$)
      END
```

```
      SUBROUTINE CEGTGN(SIGCHR,KEYCHR,JX,JY,JSEGNO,JPIKID)
CE    This subroutine decodes (x,y) reports from the terminal.
CE    For more detail, see page 5-4 of Programmers' Reference
CE    Manual for T4107/4109 Terminals.
CE    SIGCHR   ASCII code of signature character
CE    KEYCHR key character
CE    JX,JY x,y-coordinates of device position
CE    JSEGNO segment number of picked segment
CE    JPIKID pick id of segment picked
CE    The SYS$QIOW part of the routine was written by
CE    Julian Crowe, University of St Andrews Computing Laboratory.
      INTEGER STATUS,JSEGNO,JPIKID
      INTEGER JX,JY
      INTEGER*2 INPUT_CHAN,CODE
      CHARACTER*1 KEYCHR,SIGCHR
      INTEGER TIMLIMIT,
     2 INPUT_BUFF_SIZE,
     2 INPUT_SIZE
      PARAMETER (INPUT_BUFF_SIZE = 132)
      CHARACTER*132 INPUT       !BUFFER TO RECEIVE INPUT
      INCLUDE '($IODEF)'        !DEFINE SYMBOLS FOR I/O OPERATIONS
      INCLUDE '($SSDEF)'        !DEFINE STATUS VALUES
      STRUCTURE /IOSTAT_BLOCK/
      INTEGER*2 IOSTAT,
     2          TERM_OFFSET,
     2          TERMINATOR,
     2          TERM_SIZE
      END STRUCTURE             !STATUS BLOCK FOR QIOW
      RECORD /IOSTAT_BLOCK/ IOSB
      INTEGER*4 SYS$ASSIGN,
     2          SYS$QIOW
      INPUT=' '
      STATUS = SYS$ASSIGN('SYS$INPUT',    !OR USE TERMINAL NAME 'TT...'
     2          INPUT_CHAN,,)
      IF (.NOT. STATUS) CALL LIB$SIGNAL(%VAL (STATUS))
      CODE = (IO$_READVBLK)     !RETURN ON TIMEOUT IF THERE
                                !IS NO TERMINATOR
CE    TIMLIMIT = 2              !TIMEOUT LIMIT IN SECONDS: CAN BE ZERO
      STATUS = SYS$QIOW (,
     2          %VAL (INPUT_CHAN),
     2          %VAL (CODE),
     2          IOSB,
     2          ,,
     2          %REF (INPUT),
     2          %VAL (INPUT_BUFF_SIZE),
     2          ,
     2          ,,)
      IF (.NOT. STATUS)
     2                  STATUS = IOSB.IOSTAT
      IF (.NOT. STATUS)
     2                  CALL LIB$SIGNAL (%VAL (IOSB.IOSTAT))
            !DISPLAY WHAT HAS BEEN RECEIVED FROM THE TERMINAL
CE       WRITE (UNIT=*,FMT=*) 'INPUT:',INPUT (1:IOSB.TERM_OFFSET)
      CALL DECODEXY(INPUT(1:IOSB.TERM_OFFSET),SIGCHR,KEYCHR,JX,JY)
      RETURN
      END
```

```
      SUBROUTINE CEINIT
CE This sets the terminal to TEK mode,
CE enables dialog area and makes it visible.
      CHARACTER*1 ESC,RET
      ESC=CHAR(27)
      RET=CHAR(13)
      WRITE(6,60)ESC//'%!0'//RET
      WRITE(6,60)ESC//'KA1'//RET
      WRITE(6,60)ESC//'LV1'//RET
 60   FORMAT('+',A$)
      END

      SUBROUTINE CELNIN(INDEX)
CE Equivalent to LLLNIN
CE SET-LINE-INDEX
CE INDEX  line index
      CHARACTER*1 ESC,RET
      CHARACTER*4 INTSTR,CODEIN
      INTEGER INDEX
      ESC=CHAR(27)
      INTSTR=CODEIN(INDEX)
      WRITE(6,60)ESC//'ML'//INTSTR
 60   FORMAT('+',A$)
      END

      SUBROUTINE CEMOVE(IX,IY)
CE Equivalent to LLMOVE
CE MOVE
CE IX,IY coordinates to move to
      CHARACTER*1 ESC,RET
      CHARACTER*5 XYSTR,CODEXY
      INTEGER IX,IY
      ESC=CHAR(27)
      RET=CHAR(13)
      XYSTR=CODEXY(IX,IY)
      WRITE(6,60)ESC//'LF'//XYSTR//RET
 60   FORMAT('+',A$)
      END

      SUBROUTINE CEOPSG(ISEG)
CE Equivalent to LLOPSG
CE BEGIN-SEGMENT
CE ISEG segment number
      CHARACTER*1 ESC,RET
      CHARACTER*4 INTSTR,CODEIN
      INTEGER ISEG
      ESC=CHAR(27)
      INTSTR=CODEIN(ISEG)
      WRITE(6,60)ESC//'SO'//INTSTR
 60   FORMAT('+',A$)
      END

      SUBROUTINE CEPAGE
CE Equivalent to LLPAGE
CE PAGE
      CHARACTER*1 ESC,RET,FF
      ESC=CHAR(27)
      RET=CHAR(13)
      FF=CHAR(12)
```

```
      WRITE(6,60)ESC//FF//RET
 60   FORMAT('+',A$)
      END

      SUBROUTINE CEPVSG(IX,IY)
CE Equivalent to LLPVSG
CE SET-SEGMENT-PIVOT-POINT
CE IX,IY coordinates of new pivot point
      CHARACTER*1 ESC,RET
      CHARACTER*5 XYSTR,CODEXY
      INTEGER IX,IY
      ESC=CHAR(27)
      RET=CHAR(13)
      XYSTR=CODEXY(IX,IY)
      WRITE(6,60)ESC//'SP'//XYSTR//RET
 60   FORMAT('+',A$)
      END

      SUBROUTINE CERPSG(ITYPE,ISIG,ITRMSG)
CE Equivalent to LLRPSG
CE SET-REPORT-SIGNATURE-CHARS
CE ITYPE report type code or GIN device code
CE ISIG  signature character ASCII code
CE ITRMSG termination signature character ASCII code
      CHARACTER*1 ESC,RET
      CHARACTER*4 INTSTR1,INTSTR2,INTSTR3,CODEIN
      INTEGER ITYPE,ISIG,ITRMSG
      INTEGER N1,N2
      ESC=CHAR(27)
      RET=CHAR(13)
      INTSTR1=CODEIN(ITYPE)
      INTSTR2=CODEIN(ISIG)
      INTSTR3=CODEIN(ITRMSG)
      N1=2
      N2=2
 100  CONTINUE
      IF (INTSTR1(N1:N1).EQ.RET) GOTO 200
      N1=N1+1
      GOTO 100
 200  CONTINUE
      IF (INTSTR2(N2:N2).EQ.RET) GOTO 300
      N2=N2+1
      GOTO 200
 300  CONTINUE
      N1=N1-1
      N2=N2-1
      IF (ITRMSG.EQ.0) THEN
      WRITE(6,60)ESC//'IS'//INTSTR1(1:N1)//INTSTR2
      ELSE
      WRITE(6,60)ESC//'IS'//INTSTR1(1:N1)//INTSTR2(1:N2)//INTSTR3
      ENDIF
 60   FORMAT('+',A$)
      END
```

```
      SUBROUTINE CESFIL(IPAT)
CE Equivalent to LLSFIL
CE SELECT-FILL-PATTERN
CE IPAT fillpattern number
      CHARACTER*1 ESC,RET
      CHARACTER*4 INTSTR,CODEIN
      INTEGER IPAT
      ESC=CHAR(27)
      RET=CHAR(13)
      INTSTR=CODEIN(IPAT)
      WRITE(6,60)ESC//'MP'//INTSTR
 60   FORMAT('+',A$)
      END

      SUBROUTINE CESLVW(IVIEW)
CE Equivalent to LLSLVW
CE SELECT-VIEW
CE IVIEW number of view selected
      CHARACTER*1 ESC
      CHARACTER*4 INTSTR,CODEIN
      INTEGER IVIEW
      ESC=CHAR(27)
      INTSTR=CODEIN(IVIEW)
      WRITE(6,60)ESC//'RC'//INTSTR
 60   FORMAT('+',A$)
      END

      SUBROUTINE CETEXT(ILEN,TXTSTR)
CE Equivalent to LLTEXT
CE GRAPHIC-TEXT
CE ILEN device function code
CE TXTSTR text character string of length ILEN
      CHARACTER*1 ESC,RET
      CHARACTER*4 CODEIN,INTVAL
      CHARACTER*4 CILEN
      CHARACTER*70 TXTSTR
      INTEGER ILEN
      ESC=CHAR(27)
      RET=CHAR(13)
      INTVAL=CODEIN(ILEN)
CE Remove RET from end of string
      J=2
 1    CONTINUE
      IF (INTVAL(J:J).EQ.RET) GOTO 2
      J=J+1
      GOTO 1
 2    WRITE(6,60)ESC//'LT'//INTVAL(1:J-1)//TXTSTR(1:ILEN)//RET
 60   FORMAT('+',A$)
      END

      SUBROUTINE CETXIN(INDEX)
CE Equivalent to LLTXIN
CE SET-TEXT-INDEX
CE INDEX index for subsequent alpha- and graphtext
      CHARACTER*1 ESC
      CHARACTER*4 CODEIN,INTVAL
      INTEGER INDEX
      ESC=CHAR(27)
      INTVAL=CODEIN(INDEX)
```

```
      WRITE(6,60)ESC//'MT'//INTVAL
 60   FORMAT('+',A$)
      END

      SUBROUTINE CEVISG(ISEG,IVIS)
CE Equivalent to LLVISG
CE SET-SEGMENT-VISIBILITY
CE ISEG segment number
CE IVIS visibility 0 no; 1 yes
      CHARACTER*1 ESC,RET
      CHARACTER*4 INTSTR1,INTSTR2,ITEMP,CODEIN
      INTEGER ISEG,IVIS
      ESC=CHAR(27)
      RET=CHAR(13)
      INTSTR1=CODEIN(ISEG)
CE We have to get rid of the RET character at the end of this string
      J=2
 1    CONTINUE
      IF (INTSTR1(J:J).EQ.RET) GOTO 2
      J=J+1
      GOTO 1
 2    CONTINUE
      INTSTR2=CODEIN(IVIS)
      WRITE(6,60)ESC//'SV'//INTSTR1(1:J-1)//INTSTR2
 60   FORMAT('+',A$)
      END

      SUBROUTINE CEVWPT(ILRX,ILRY,IUPX,IUPY)
CE Equivalent to LLVWPT
CE SET-VIEWPORT
CE ILRX lower left x coordinate
CE ILRY lower left y coordinate
CE IUPX upper right x coordinate
CE IUPY upper right y coordinate
      CHARACTER*1 ESC,RET
      CHARACTER*5 XYSTR1,XYSTR2,CODEXY
      INTEGER ILRX,ILRY,IUPX,IUPY
      ESC=CHAR(27)
      RET=CHAR(13)
      XYSTR1=CODEXY(ILRX,ILRY)
      XYSTR2=CODEXY(IUPX,IUPY)
      WRITE(6,60)ESC//'RV'//XYSTR1//XYSTR2//RET
 60   FORMAT('+',A$)
      END

      SUBROUTINE CEWIND(ILRX,ILRY,IUPX,IUPY)
CE Equivalent to LLWIND
CE SET-WINDOW
CE ILRX lower left x coordinate
CE ILRY lower left y coordinate
CE IUPX upper right x coordinate
CE IUPY upper right y coordinate
      CHARACTER*1 ESC,RET
      CHARACTER*5 XYSTR1,XYSTR2,CODEXY
      INTEGER ILRX,ILRY,IUPX,IUPY
      ESC=CHAR(27)
      RET=CHAR(13)
      XYSTR1=CODEXY(ILRX,ILRY)
      XYSTR2=CODEXY(IUPX,IUPY)
```

```
      WRITE(6,60)ESC//'RW'//XYSTR1//XYSTR2//RET
 60   FORMAT('+',A$)
      END

      CHARACTER*4 FUNCTION CODEIN(INTVAL)
CE This function encodes integers into host syntax.
CE For more details, see page 5-2 of the 4107/4109
CE Programmers' Reference Manual.
      CHARACTER*4 CODE
      CHARACTER*1 RET
      INTEGER INTVAL,INTEMP,ICOUNT,HII1,HII2,LOI
      LOGICAL POS
      DATA RET / 13 /
      ICOUNT=1
      CODE=' '
      POS=.TRUE.
      IF (INTVAL.LT.0) POS=.FALSE.
      INTEMP=ABS(INTVAL)
      HII1=INTEMP/1024
      IF (HII1.GT.0) THEN
      CODE(ICOUNT:ICOUNT)=CHAR(64+HII1)
      ICOUNT=ICOUNT+1
      ENDIF
      HII2= MOD(INTEMP,1024)/16
      IF (HII2.GT.0 .OR. HII1.GT.0) THEN
      CODE(ICOUNT:ICOUNT)=CHAR(64+HII2)
      ICOUNT=ICOUNT+1
      ENDIF
      LOI=MOD(INTEMP,16)
      IF (POS) LOI=LOI+16
      CODE(ICOUNT:ICOUNT)=CHAR(32+LOI)
      ICOUNT=ICOUNT+1
      CODE(ICOUNT:ICOUNT)=RET
      CODEIN=CODE(1:ICOUNT+1)
      RETURN
      END

      CHARACTER*5 FUNCTION CODEXY(INX,INY)
CE  This function encodes (x,y) values into host syntax
CE  For more detail, see page 5-4 of Programmers' Reference
CE  Manual for T4107/4109 Terminals.
      INTEGER INX,INY
      INTEGER HIX,LOX,HIY,LOY,EXTRAX,EXTRAY,EXTRA
      HIX=32+(INX/128)
      HIY=32+(INY/128)
      LOX=64+(MOD(INX,128)/4)
      LOY=96+(MOD(INY,128)/4)
      EXTRAX=MOD(INX,4)
      EXTRAY=MOD(INY,4)
      EXTRA=96+(EXTRAY*4)+EXTRAX
      CODEXY=CHAR(HIY)//CHAR(EXTRA)//CHAR(LOY)//CHAR(HIX)//CHAR(LOX)
      RETURN
      END
```

```
      SUBROUTINE DECODEXY(CODEXY,SIGCHR,KEYCHR,INX,INY)
CE    This subroutine decodes (x,y) reports from the terminal.
CE    For more detail, see page 5-4 of Programmers' Reference
CE    Manual for T4107/4109 Terminals.
      CHARACTER*7 CODEXY
      CHARACTER*1 KEYCHR,SIGCHR
      INTEGER HIX,HIY,LOX,LOY,EXTRA,EXTRAX,EXTRAY
      SIGCHR=CODEXY(1:1)
      KEYCHR=CODEXY(2:2)
      HIY=ICHAR(CODEXY(3:3))
      EXTRA=ICHAR(CODEXY(4:4))
      LOY=ICHAR(CODEXY(5:5))
      HIX=ICHAR(CODEXY(6:6))
      LOX=ICHAR(CODEXY(7:7))
      HIY=HIY-32
      HIY=HIY*128
      EXTRA=EXTRA-32
      EXTRAX=MOD(EXTRA,4)
      EXTRAY=MOD(EXTRA,16)/4
      LOY=LOY-32
      LOY=LOY*4
      HIX=HIX-32
      HIX=HIX*128
      LOX=LOX-32
      LOX=LOX*4
      INX=HIX+LOX+EXTRAX
      INY=HIY+LOY+EXTRAY
      RETURN
      END
```

REFERENCES

1) J.C. Holland, in "Cancer Medicine", J.F. Holland, E. Frei III, (Eds), Lea & Febiger, Philadelphia (1982)

2) H.C. Pitot, in "Cancer, Principles And Practice Of Oncology", V.T. DeVita Jr., S. Hellman, S.A. Rosenberg, (Eds), J.B. Lippincott, Philadelphia (1985)

3) C. Thomson, accepted for J. Quant. Chem (Quant. Biol. Symp.) QB13 (1986)

4) C. Thomson, accepted for Acc. Chem. Res.

5) R. Doll, R. Peto, "The Causes Of Cancer", Oxford University Press, Oxford (1981)

6) F.F. Becker, in "Cancer, A Comprehensive Treatise, Vol I: Etiology: Chemical And Physical Carcinogenesis", F.F. Becker, (Ed), Plenum Press, New York (1982)

7) P. Buell, J.E. Dunn, Cancer 18 (1965) 656, quoted by J. Cairns, "Cancer: Science And Society", W.H. Freeman, San Francisco (1978)

8) "Diet, Nutrition and Cancer", Assembly Of Life Sciences, National Research Council Report, National Academy Press, Washington DC (1982), quoted in G.R. Newell, in "Cancer, Principles And Practice Of Oncology", V.T. DeVita Jr., S. Hellman, S.A. Rosenberg, (Eds), J.B. Lippincott, Philadelphia (1985)

9) M.S.C. Birbeck, in "Scientific Foundations Of Oncology", T. Symington, R.L. Carter, (Eds), William Heinemann Medical Books, London (1976)

10) M.F.A. Woodruff, "The Interaction Of Cancer And Host: Its Therapeutic Significance", Grune & Stratton, New York (1980)

11) H.C. Pitot, "Fundamentals Of Oncology", Marcel Dekker, New York (1981)

12) T.J. McElwain, in "Scientific Foundations Of Oncology", T.Symington, R.L. Carter, (Eds), William Heinemann Medical Books, London (1976)

13) S.E. Jones, S.E. Salmon, "Adjuvant Therapy Of Cancer II', Grune & Stratton, New York (1979)

14) A. Szent-Gyorgyi, "The Living State And Cancer", Marcel Dekker, New York (1978)

15) L. Pauling, Chem. Eng. News, 24 (1946) 1375

16) R.A. Wallace, A.N. Kurtz, C. Niemann, Biochemistry, 2 (1963) 824

17) H. Lineweaver, D. Burk, J. Amer. Chem. Soc., 56 (1934) 658

18) M. Dixon, E.C. Webb, "Enzymes", Longman, London (1979)

19) N.K. Ghosh, W.H. Fishmann, J. Biol. Chem., 241 (1966) 2516

20) J. Fastrez, A.R. Fersht, Biochemistry, 12 (1973) 1067

21) G.B. Fitzgerald, M.M. Wick, Biochem. Pharmac., 34 (1985) 353

22) T.G. O'Brien, Cancer Res., 36 (1976) 2644

23) F. Marks, S. Bertsch, G. Furstenberger, Cancer Res., 39 (1979) 4183

24) J. Bartholeyns, P. Mamont, P. Casara, Cancer Res., 44 (1984) 4972

25) C. Danzin, P. Casara, N. Claverie, B.W. Metcalf, M.J. Jung, Biochem. Biophys. Res. Commun., 116 (1983) 237

26) B.W. Metcalf, P. Bey, C. Danzin, M.J. Jung, P. Casara, J.P. Vevert, J. Amer. Chem. Soc., 100 (1978) 2551

27) W. Troll, A. Klassen, A. Janoff, Science, 169 (1970) 1211

28) E.L. Smith, in "The Enzymes, Chemistry And Mechanism Of Action", J.B. Sumner, K. Myrback, (Eds), Academic Press, New York (1951)

29) M. Hozumi, M. Ogawa, T. Sugimura, T. Takeuchi, H. Umezawa, Cancer Res., 32 (1972) 1725

30) T. Matushima, T. Kaziko, T. Kawachi, K. Hara, T. Sugimura, T. Takeuchi, H. Umezawa, in "Fundamentals In Cancer Prevention", P.N. Magee et al., (Eds), University Of Tokyo Press, Tokyo (1975)

31) A.R. Kennedy, J.B. Little, Nature, 276 (1978) 825

32) T. Kuroki, C. Drevon, Cancer Res., 39 (1979) 2755

33) J.E. Cleaver, M.J. Banda, W. Troll, C. Borek, Carcinogenesis, 7 (1986) 323

34) S.B. Bhattacharjee, N. Bhattacharyya, Carcinogenesis, 7 (1986) 1267

35) T.J. Slaga, in "Modifiers Of Chemical Carcinogenesis: An Approach To The Biochemical Mechanism And Cancer Prevention, Vol V", Raven Press, New York (1980)

36) S.M. Fischer, G.L. Gleason, G.D. Mills, T.J. Slaga, Cancer Lett., 10 (1980) 343

37) S.M. Fischer, G.D. Mills, T.J. Slaga, Carcinogenesis, 3 (1982) 1243

38) H. Uzumaki, S. Yamamoto, R. Kato, Carcinogenesis, 7 (1986) 289

39) A. Dipple, R.C. Moschel, C.A.H. Bigger, in "Chemical Carcinogens", L.E. Searle, (Ed), ACS Monograph 182, ACS, Washington DC (1984)

40) G.T. Bowden, T.J. Slaga, B.G. Shapas, R.K. Boutwell, Cancer Res., 34 (1974) 2634

41) H.V. Gelboin, F. Wiebel, L. Diamond, Science, 170 (1970) 169

42) C.M. Ireland, A. Eastman, E. Bresnick, Carcinogenesis 5 (1984) 187

43) R. Vince, W.B. Wadd, Biochem. Biophys. Res. Commun., 35 (1969) 593

44) A. Szent-Gyorgyi, "Electronic Biology And Cancer", Marcel Dekker, New York (1976)

45) A-C. Aronsson, B. Mannervik, Biochem. J., 165 (1977) 503

46) M. Brandt, "The Chemical Mechanism Of Glyoxalase I", unpublished report, Laboratory of Dr R. Brandt, Medical College Of Virginia, Richmond, Virginia.

47) B. Mannervik, in "Enzymatic Basis Of Detoxication, Vol II", Academic Press (1980)

48) F. Jordan, J.F. Cohen, C-T. Wang, J.M. Wilmott, S.S. Hall, D.C. Foxall, Drug Metab. Rev., 14 (1983) 723

49) K.T. Douglas, S. Shinkai, Angew. Chem. Int. Ed. Engl., 24 (1985) 31

50) H.D. Dakin, H.W. Dudley, J. Biol. Chem., 14 (1913) 155

51) C. Neuberg, Biochem. Z., 49 (1913) 202

52) E. Gillespie, Nature, 277 (1979) 135

53) E. Gillespie, J. Immunol., 121 (1978) 923

54) W.H. Elliott, Biochim. Biophys. Acta, 29 (1958) 446

55) G. Urata, S. Granick, J. Biol. Chem., 238 (1963) 811

56) L. Egyud, A Szent-Gyorgyi, Proc. Nat. Acad. Sci. USA, 55 (1965) 388

57) A. Szent-Gyorgyi, Science, 161 (1968) 988

58) E. Marmstal, A-C. Aronsson, B. Mannervik, Biochem. J., 183 (1979) 23

59) B. Mannervik, T. Bartfai, B. Gorna-Hall, J. Biol. Chem., 249 (1974) 901

60) B. Mannervik, B. Gorna-Hall, T. Bartfai, Eur. J. Biochem., 37 (1973) 270

61) T. Bartfai, K. Ekwall, B. Mannervik, Biochemistry, 12 (1973) 387

62) K.A. Davis, G.R. Williams, Biochim. Biophys. Acta, 113 (1966) 393

63) A-C. Aronsson, E. Marmstal, B. Mannervik, Biochem. Biophys. Res. Commun., 81 (1978) 1235

64) A-C. Aronsson, S. Sellin, G. Tibbelin, B. Mannervik, Biochem. J., 197 (1981) 67

65) K. Ekwall, B. Mannervik, Arch. Biochem. Biophys. 137 (1970) 128

66) B. Mannervik, E. Marmstal, K. Ekwall, B. Gorna-Hall, Eur. J. Biochem., 53 (1975) 327

67) S. Sellin, A-C. Aronsson, B. Mannervik, L.E.G. Eriksson, Acta Chemica Scand., B35 (1981) 229

68) S. Sellin, L.E.G. Eriksson, B. Mannervik, Biochemistry, 21 (1982) 4850

69) L. Garcia-Iniguez, L. Powers, B. Chance, S. Sellin, B. Mannervik, A.S. Mildvan, Biochemistry, 23 (1984) 685

70) P. Rosevear, S. Sellin, B. Mannervik, I.D. Kuntz, A.S. Mildvan, J. Biol. Chem., 259 (1984) 11436

71) P.R. Laurence, C. Thomson, Theoret. Chim. Acta, 57 (1980) 25

72) S. Sellin, P.R. Rosevear, B. Mannervik, A.S. Mildvan, J. Biol. Chem., 257 (1982) 10023

73) P.R. Rosevear, R.V.J. Chari, J.W. Kozarich, S. Sellin, B. Mannervik, A.S. Mildvan, J. Biol. Chem., 258 (1983) 6823

74) C.E.F. Griffis, L.H. Ong, L. Buettner, D.J. Creighton, Biochemistry, 22 (1983) 2945

75) E. Racker, J. Biol. Chem., 190 (1951) 685

76) V. Franzen, Chem. Ber., 89 (1956) 1020

77) V. Franzen, Chem. Ber., 90 (1957) 623

78) I.A. Rose, Biochim. Biophys. Acta, 25 (1957) 214

79) J.F. Hine, G.F. Koser, J. Org. Chem., 36 (1971) 3591

80) D.L. Vander Jagt, L-P. Han, C.H. Lehman, J. Org. Chem., 37 (1972) 4100

81) S.S. Hall, A.M. Doweyko, F. Jordan, J. Amer. Chem. Soc., 100 (1978) 5934

82) S.S. Hall, A. Poet, Tet. Lett., (1970) 2867

83) T. Okuyama, K. Kimura, T. Fueno, Bull. Chem. Soc. Jpn, 55 (1982) 1493

84) S. Shinkai, T. Ide, O. Manabe, Chem. Lett. (1978) 583

85) S. Shinkai, T. Yamashita, O. Manabe, J. Chem. Soc. Chem. Commun., (1979) 301

86) S. Shinkai, T. Yamashita, Y. Kusano, T. Ide, O. Manabe, J. Amer. Chem. Soc., 102 (1980) 2335

87) P. Hemmerich, V. Massey, H. Fenner, FEBS Lett., 84 (1977) 5

88) S. Shinkai, T. Yamashita, Y. Kusano, O. Manabe, J. Amer. Chem. Soc., 103 (1981) 2070

89) K. Ueda, S. Shinkai, K.T. Douglas, J. Chem. Soc. Chem. Commun., (1984) 371

90) K.T. Douglas, A.J. Quilter, S. Shinkai, K. Ueda, Biochim. Biophys. Acta, 829 (1985) 119

91) D.L. Vander Jagt, L-P.B. Han, C. Lehman, Biochemistry, 11 (1972) 3735

92) R. Vince, S. Daluge, W.B. Wadd, J. Med. Chem., 14 (1971) 402

93) W.O. Kermack, N.A. Matheson, Biochem. J., 65 (1957) 48

94) E.H. Cliffe, S.G. Waley, Biochem. J., 79 (1960) 475

95) M.V. Kester, J.A. Reese, S.J. Norton, J. Med. Chem., 17 (1974) 413

96) R. Vince, S. Daluge, J. Med. Chem., 14 (1971) 35

97) P.A. Lyon, R. Vince, J.Med Chem., 20 (1977) 77

98) B. Oray, S.J. Norton, Biochem. Biophys. Res. Commun., 95 (1980) 624

99) H. Chimura, H. Nakamura, T. Takita, H. Umezawa, K. Kato, S. Saito, T. Tomisawa, Y. Iitaka, J. Antibiot., 28 (1975) 743

100) H. Nakamura, Y. Iitaka, S. Kurasawa, T. Takeuchi, H. Umezawa, Agr. Biol. Chem., 40 (1976) 1781

101) M. Iio, K. Okabe, H. Omura, J. Nutr. Sci. Vitaminol., 22 (1976) 53

102) K.T. Douglas, I.N. Nadvi, FEBS Lett., 106 (1979) 393

103) K.T. Douglas, J. Ghotb-Sharif, Biochim. Biophys. Acta, 748 (1983) 184

104) R.B. Brandt, M.E. Brandt, M.E. April, C.Thomson, Int. J. Quant. Chem. (Quant. Biol. Symp.) QB9 (1982) 335

105) C. Thomson, R.B. Brandt, Int. J. Quant. Chem. (Quant. Biol. Symp.) QB10 (1983) 357

106) R.B. Brandt, J. Laux, C. Thomson, M.A. Johnson, M. Gross, Int. J. Quant. Chem. (Quant. Biol. Symp.) QB11 (1984) 195

107) R.B. Brandt, J.E. Laux, S.W. Yates, M.R. Boots, C. Thomson, C. Edge, accepted for Int. J. Quant. Chem. (Quant. Biol. Symp.) QB13 (1986)

108) C. Hansch, Acc. Chem. Res., 3 (1969) 232

109) ChemGraf, Chemical Design Ltd, Oxford

110) P. Quarendon, C.B. Naylor, W.G. Richards, J. Mol. Graphics, 2 (1984) 4

111) H. Nakamura, M. Kusunoki, N. Yasuoka, J. Mol. Graphics, 2 (1984) 14

112) A. Szabo, N.S. Ostlund, "Modern Quantum Chemistry: Introduction To Advanced Electronic Structure Theory", Macmillan, New York (1982)

113) E. Steiner, "The Determination And Interpretation Of Molecular Wave Functions", Cambridge University Press, Cambridge (1976)

114) W.J. Hehre, R.F. Stewart, J.A. Pople, J. Chem. Phys., 51 (1969) 2657

115) W.J. Hehre, R. Ditchfield, R.F. Stewart, J.A. Pople, J. Chem. Phys., 52 (1970) 2769

116) J.S. Binkley, J.A. Pople, W.J. Hehre, J. Amer. Chem. Soc. 102 (1980) 939

117) M.S. Gordon, J.S. Binkley, J.A. Pople, W.J. Pietro, W.J. Hehre, J. Amer. Chem. Soc., 104 (1982) 2797

118) P.C. Hariharan, J.A. Pople, Theoret. Chim. Acta 28 (1973) 213

119) W.J. Hehre, J. Amer. Chem. Soc., 97 (1975) 5308

120) W.J. Pietro, M.M. Francl, W.J. Hehre, D.J. DeFrees, J.A. Pople, J.S. Binkley, J. Amer. Chem. Soc., 104 (1982) 5039

121) E. Scrocco, J. Tomasi, Top. Curr. Chem., 42 (1973) 95

122) D.G. Truhlar in "Chemical Applications Of Atomic And Molecular Electrostatic Potentials", P. Politzer, D.G. Truhlar, (Eds), Plenum Press, New York (1981)

123) H. Weinstein, R. Osman, J.P. Green, S. Topiol in "Chemical Applications Of Atomic And Molecular Electrostatic Potentials", P. Politzer, D.G. Truhlar, (Eds), Plenum Press, New York (1981)

124) H. Weinstein, S. Maayani, S. Srebenik, S. Cohen, M. Sokolovsky, Mol. Pharmacol., 11 (1975) 671

125) R. Osman, H. Weinstein, Chem. Phys. Lett., 49 (1977) 69

126) W.D. Edwards, H. Weinstein, Chem. Phys. Lett., 56 (1978) 582

127) K. Kitaura, K. Morokuma, Int. J. Quant. Chem., 10 (1976) 325

128) P. Pulay, in "Applications Of Electronic Structure Theory, Vol 4", H.F. Schaefer, (Ed), Plenum Press, New York (1977)

129) J.S. Binkley, R.A. Whiteside, R. Krishnan, R. Seeger, D.J. De Frees, H.B. Schlegel, S. Topiol, L.R. Kahn, J.A. Pople, QCPE No 13 (1981) 406

130) M. Frisch, GAUSSIAN 82 Release A, Carnegie-Mellon University (1983)

131) R. Fletcher, M.J.D. Powell, Compt. J., 6 (1963) 163

132) H.B. Schlegel, J. Comp. Chem., 3 (1982) 214

133) B.A. Murtaugh, R.W.H. Sargent, Compt. J. 13 (1980) 185

134) H.B. Schlegel, Theoret. Chim. Acta, 66 (1984) 333

135) C. Thomson, unpublished work

136) L-P.B. Han, L.M. Davison, D.L. Vander Jagt, Biochim. Biophys. Acta, 445 (1976) 486

137) J.J. Kaufman, P.C. Hariharan, F.C. Tobin, C. Petrongolo in "Chemical Applications Of Atomic And Molecular Electrostatic Potentials", P. Politzer, D.G. Truhlar, (Eds), Plenum Press, New York (1981)

138) J.R. Ball, unpublished work

139) D. Peters, M. Sana, QCPE 360 (1978) DENPOT

140) J.D. Foley, A. Van Dam, "Fundamentals Of Interactive Computer Graphics", Addison-Wesley, Reading, Massachusetts (1982)

141) M. Connolly, QCPE 429 MS

142) M. Connolly, J. Amer. Chem. Soc., 107 (1985) 1118

143) U.C. Singh, P. Kollman, QCPE Bulletin 2 (1982) 117 QCPE 446

144) W.M. Newman, R.F. Sproull, "Principles Of Interactive Computer Graphics" McGraw Hill, Tokyo (1984)

145) D.F. Rogers, "Procedural Elements For Computer Graphics", McGraw-Hill, Singapore (1985)

146) "4107/4109 Computer Display Terminals: Programmer's Reference Manual", Tektronix, Beaverton, Oregon (1983)

147) "VAX/VMS System Services Reference Manual", DEC, Maynard, Massachusetts (1980)

148) J. Crowe, personal communication.

149) R. Lavery, B. Pullman, Int. J. Quant. Chem. (Quant. Biol. Symp.) QB6 (1979) 467

150) H. Tatewaki, S. Huzinaga, J. Comput. Chem., 1 (1980) 205

151) S. Oae, W. Tagaki, A. Ohno, J. Amer. Chem. Soc., 83 (1961) 5036

152) S. Oae, A. Ohno, W. Tagaki, Chem. & Ind., (1962) 304

153) S. Oae, W. Tagaki, A. Ohno, Tetrahedron, 20 (1964) 427

154) S. Oae, W. Tagaki, A. Ohno, Tetrahedron, 20 (1964) 437

155) A. Streitwieser, J.E. Williams, J. Amer. Chem. Soc., 97 (1975) 191

156) S. Wolfe, A. Rauk, I.G. Csizmadia, J. Amer. Chem. Soc., 89 (1967) 5710

157) A. Rauk, I.G. Csizmadia, Can. J. Chem., 46 (1968) 1205

158) S. Wolfe, L.A. LaJohn, F. Bernardi, A. Mangini, G. Tonachini, Tet. Lett., 24 (1983) 3789

159) W.J. Hehre, R. Ditchfield, J.A. Pople, J. Chem. Phys. 56 (1972) 2257

160) R. Wolfenden, Nature, 223 (1969) 704

161) S. Bell, J.S. Crighton, J. Chem. Phys., 80 (1984) 2464

162) C.A. Reynolds, Ph.D. Thesis, St Andrews University (1985)

163) J. Gripenberg in "The Chemistry Of Flavonoid Compounds", T.A. Geissman (Ed), Pergamon Press, London (1962)

164) M.L. Vickery, B. Vickery, "Secondary Plant Metabolism", MacMillan, London (1981)

165) J. Okuda, I. Miwa, K. Inagaki, T. Horie, M. Nakayama, Biochem. Pharmacol., 31 (1982) 3807

166) J. Okuda, I. Miwa, K. Inagaki, T. Horie, M. Nakayama, Chem. Pharm, Bull., 32 (1984) 767

167) E.L. Wheeler, D.L. Berry, Carcinogenesis, 7 (1986) 33

168) R. Kato, T. Nakadate, S. Yamamoto, T. Sugimura, Carcinogenesis, 4 (1983) 1301

169) J.T. Kellis Jr., L.E. Vickery, Science, 225 1032

170) M. Gschwendt, F. Horn, W. Kittstein, F. Marks, Biochem. Biophys. Res. Commun., 117 (1983) 444

171) A.K. Srivastava, Biochem. Biophys. Res. Commun., 131 (1985) 1

172) Y. Tsuchiya, M. Shimizu, Y. Hiyama, K. Itoh, Y. Hashimoto, M. Nakayama, T. Horie, N. Morita, Chem. Pharm. Bull., 33 (1985) 3881

173) R.L. Chang, M-T. Huang, A.W. Wood, C-Q. Wong, H-L. Newmark, H. Yagi, J.M. Sayer, D.M. Jerina, A.H. Conney, Carcinogenesis 6 (1985) 1127

174) M-T. Huang, A.W. Wood, H.L. Newmark, J.M. Sayer, H. Yagi, D.M. Jerina, A.H. Conney, Carcinogenesis, 4 (1983) 1631

175) W.L. Alworth, T.J. Slaga, Carcinogenesis, 6 (1985) 487

176) C. Thomson, C. Edge, J. Mol. Struct. (Theochem), 121 (1985) 173

177) S. Diner, J.P. Malrieu, F. Jordan, M. Gilbert, Theoret. Chim. Acta, 15 (1969) 100

178) P.R. Laurence, Ph.D. Thesis, St Andrews University (1982)

179) R.D.H. Murray, J. Mendez, S.A. Brown, "The Natural Coumarins: Occurrence Chemistry And Biochemistry", John Wiley & Sons, Chichester (1982)

180) K. Sekiya, H. Okuda, S. Arichi, Biochim. Biophys. Acta, 713 (1982) 68

181) B.L. Van Duuren, B.M. Goldschmidt, J. Nat. Cancer Inst., 56 (1976) 1237

182) T. Ohta, K. Watanabe, M. Moriya, Y. Shirasu, T. Kada, Mutation Res., 117 (1983) 135

183) T. Ohta, K. Watanabe, M. Moriya, Y. Shirasu, T. Kada, Mutation Res., 107 (1983) 219

184) W.S. Woo, K.H. Shin, C.K. Lee, Biochem. Pharmacol., 32 (1983) 1800